

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 22:21:07 ; Search time 1836.85 seconds

(without alignments)  
5707.703 Million cell updates/sec

Title:  
Perfect score:

US-09-484-577a-3  
501

Sequence:  
1 actctccagcctctcaccga.....cagaatgcgcgatgatcat 501

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

al number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Genembl:\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_ov:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_hgt\_hum:\*  
31: em\_hgt\_inv:\*  
32: em\_hgt\_other:\*  
33: em\_hgtgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

C	1	92.4	18.4	3526	1	AF273216	AF273216 Rhizobium
C	2	88.8	17.7	28804	1	SS051197	US1197 Spingomonas
C	3	88.8	17.7	28804	6	AR068625	AR068625 Sequence
C	4	86.8	17.3	15552	1	AE004049	AE004049 Xylella f
C	5	86	17.2	1428	1	NME391263	AF121772 Neisseria
C	6	86	17.2	4254	1	AF121772	AF121772 Neisseria
C	7	86	17.2	7824	1	NME391260	AE002524 Neisseria
C	8	86	17.2	11381	1	AE002524	AE002524 Neisseria
C	9	86	17.2	326301	1	NMA622491	AX044033 Sequence
C	10	86	17.2	349980	6	AX044033	AX044033 Sequence
C	11	86	17.2	349980	6	AX044034	AL646077 Neisseria
C	12	80.4	16.0	189050	1	AL646077	AL646077 Neisseria
C	13	69.8	13.9	334050	1	AJ414151	AF193064 Caulobact
C	14	62.6	12.5	6645	1	BPCYADE	AF062345 Caulobact
C	15	56.8	11.3	7493	1	AF193064	AF193064 Caulobact
C	16	56.8	11.3	13759	1	AE005779	AE005779 Caulobact
C	17	56.8	11.3	19883	1	AF062345	AF062345 Caulobact
C	18	54.2	10.8	15441	1	SC011	AL646083 Streptomy
C	19	53.2	10.6	208050	1	AL646083	AL646083 Streptomy
C	20	51.6	10.3	5120	1	APCLYIAD	XE1112 A. pleuropne
C	21	51.6	10.3	8292	1	APCLYIAD	XE1112 A. pleuropne
C	22	51.6	10.3	8370	6	AR091872	AR091872 Sequence
C	23	48.6	9.7	348411	1	ACNAPAPAB	ME5808 Actinobacti
C	24	48.4	9.7	4247	1	ACNAPAPAB	ME5808 Actinobacti
C	25	48	9.6	7413	1	PAAPRAPG	AE004554 Pseudomon
C	26	48	9.6	11588	1	AE004554	AE004554 Pseudomon
C	27	47.8	9.5	6400	6	E30060	E30060 ABC Transpo
C	28	47.8	9.5	11273	1	AB032389	AB032389 Pseudomon
C	29	47.6	9.5	11383	1	AE004761	AE004761 Pseudomon
C	30	47.2	9.4	299350	1	SMES91786	AL591786 Sinorhizo
C	31	47	9.4	18857	1	AE006947	AE006947 Mycobacte
C	32	47	9.4	35420	1	MTCY22610	284724 Mycobacte
C	33	47	9.4	42741	1	MSGY423	AD000014 Mycobacte
C	34	46.6	9.3	333500	1	AP003590	AP003590 Nostoc sp
C	35	46.4	9.3	10029	1	AE009262	AE009262 Agrobacte
C	36	46.4	9.3	10235	1	AE008350	AE008350 Agrobacte
C	37	45.6	9.1	10072	1	AE005876	AE005876 Agrobacte
C	38	45.6	9.1	340857	1	AP003010	AP003010 Mesorhizo
C	39	45.2	9.0	215050	1	AL646057	AL646057 Ralstonia
C	40	45.2	9.0	264050	1	AL627279	AL627279 Ralstonia
C	41	45.2	9.0	274050	1	AL627279	AL627279 Ralstonia
C	42	45.2	9.0	333800	1	SMES91792	AL591792 Sinorhizo
C	43	44.8	8.9	22991	1	SCD19	AL592149 Streptomy
C	44	44.4	8.9	3368	1	S67013	S67013 prtd-protea
C	45	44.4	8.9	4561	1	SMASADE	X81195 S.marcescen

## ALIGNMENTS

RESULT	1	AF273216	3526 bp	DNA	Linear	BCT 17-JUL-2001
LOCUS		AF273216				
DEFINITION		Rhizobium leguminosarum bv. viciae bacteriophages (trza) gene, partial cds; ABC transporter Rcbp (rzbp) gene, complete cds; and transporter component Rcbp (rzbp) gene, partial cds.				
ACCESSION		AF273216				
VERSION		AF273216.1				
KEYWORDS		GI:10952793				
SOURCE						
ORGANISM		Rhizobium leguminosarum bv. viciae.				
		Rhizobium leguminosarum bv. viciae				
		Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;				
REFERENCE		1 (bases 1 to 3526)				
AUTHORS		Venter, A.P., Twelker, S., Oresnik, I.J. and Hynes, M.F.				
TITLE		Analysis of the genetic region encoding a novel rhizobiocin from Rhizobium leguminosarum bv. viciae strain 306				
JOURNAL		Rhizobium journal of microbiology. 47 (6), 495-502 (2001)				
MEDLINE		21360352				
PUBMED		11467725				
REFERENCE		2 (bases 1 to 3526)				
AUTHORS		Venter, A.P., Twelker, S. and Hynes, M.F.				
TITLE		Direct Submission				



gene  
CDS  
/protein\_id="AAC44059.1"  
/db\_xref="GI:1314566"  
/translation="MKKLVLTVAANAALAVSGGSGKEGLDKGVAVSDGDEITIF  
ELNAEVAOTVPGQTDRIKLAEQMALQRIIRKLLAKYARFQKIDKTPVPIQORAD  
LITLTLNRKIAAGIAQPTDAIEAIOYFAAHPDFEAFQKIKYISIDKIVPELQORAD  
FAPKLTIDLOAKLITADVQFRAFPRTQDPAALPEPLAKIALPAQELFILTQGV  
TANVTGTTIOTPLAGDARLANTLTTERFSAAADQALNDRKAKESKTYQPGYSP  
PPOJGGAAPAAPAPAPAPNAAQ"  
7588. .8646  
/gene="spsk"  
7588. .8646  
/gene="spsk"  
/function="sphingan polysaccharide synthesis"  
/codon\_start=1  
/transl\_table=11  
/evidence=experimental  
/product="glycosyl transferase"  
/protein\_id="AAC44060.1"  
/db\_xref="GI:1314567"  
/translation="MGDGAETVTEAKAGKPLKMCILAAAGGHLRDLLESYKREH  
DYFFVTEDLALGRSLAEKHSVALVHYVIALGQAKLGHPIRMAGARINLRQSLSTIRK  
KFDVVISTAGAVYFTALALAKSGAKFYHIESEAFEDHSPAFGKWKICIAVTVYQSA  
AKQWIPSAEGLDTPFRLLDTPPPKQALFPAVGTGLEPRLYQAVDLAKAGGIPCK  
LVLQDQDLDADPGIPIDEIIRITFPEDLOLIRPADAVTCHGGTGSIVLALRGCRV  
VAFPRRHIDGHEHDHDEIEIQTFFADRLLIHAVDERELGAVEAKATEPQLATTDH  
TLAQRLEHLLEAOWAKR"  
8643. .9509  
/gene="spst"  
8643. .9509  
/gene="spst"  
/function="sphingan polysaccharide synthesis"  
/codon\_start=1  
/transl\_table=11  
/evidence=experimental  
/product="glycosyl transferase"  
/protein\_id="AAC44061.1"  
/db\_xref="GI:1314568"  
/translation="MSAPRISVVIPIHYNDPSLRQCIDLQHOTIGREAFETIVGNN  
SPGCLAVEAAGARAVITLLEKGAAPARNGAAAEQGEILFTSDCVCEGWLGA  
GVAHVAGFRGVGHMIVLKEPGLTGAELMALPDNEGSEYRAKPTVYANLFTVRA  
DFEHWGGRFVGEDELMCHCRATGALIDAYEAASVNGEPRPDATLVKTRIORE  
LLEFNIRPGRIRFWLARSITLOPALIPADTAKLIRTPGTRGRIAAVGLVRLRFRA  
GAGLQLLGKPI"  
compement(9552. .10940)  
/gene="spst"  
compement(9552. .10940)  
/gene="spst"  
/function="secretion of sphingan polysaccharide"  
/codon\_start=1  
/transl\_table=11  
/evidence=experimental  
/product="unknown"  
/protein\_id="AAC44062.1"  
/db\_xref="GI:1314569"  
/translation="MAVSGLKELMPEGRNDVPEEDGYEPLIATVVPHRDAHAGRP  
DHPTEASALDRLDRRREREITRARALATPTTQPVADRSSPAGRELEVLRLL  
SAIESQRSHVYIGERGIKTSLLVILNDVARESSYIVSYATCGNANFADLFAVLE  
VYDLPFHRQSVNAGEAESGSLADRLPGSDSELDLDCADITGTAVLLIIDEYDR  
FEVEENMIALGTEASGLRPDPNLTIRIHLMLAGSYFARLLCHNALEALDQGLTVD  
IGHERRAIDALITELGKMSPAVLEMKRFVGGRYDPLVAALGASRSIDGMFTGOAV  
VELLGGSTTAAQVQELDALSRQLGLESDAOGERRRFNDLSLPLYLMLVGRLLI  
DSGTEERLAAV"  
11369. .12867  
/gene="spst"  
11369. .12867  
/gene="spst"  
/function="sphingan polysaccharide synthesis"  
/codon\_start=1  
/transl\_table=11  
/evidence=experimental  
/product="unknown"  
/protein\_id="AAC44063.1"





```

/note="similar to SP|P10089 (percent identity: 61 %/query
alignment coverage: 96.9 %/subject alignment coverage:
98.7 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder/Start codon shift: 21"
/codon_start=1
/transl_table=11
/product="toxin secretion ABC transporter ATP-binding
protein"
/protein_id="AA085196.1"
/db_xref="GI:9107580"
/translation="MTYKSPVLTQNGASNPQDSALHGLVLAQPHGIADPPIA
HGRSGELFDETLTLLAAKGLKAKVYQPTVRIAMALPALIPGTAIVYK
YKRLIHLADQRPRTISHDEFTARAGRLQVAVASVIGLAKPDSWPIPAYK
YKRLAEVAVSLFIQIFALVTPLEFYQVMDKLVQVQALSTLEVIYGLISTLAEVL
LTLTFAVHTSKIDDELGARGRPHLLALPLAYESRVDGTARVLELNIHREL
TGOALTSVLDLFTVYVLSVMSYSSMLLVLSLPLASLTTPVLRKRLDERF
ARGADNOALFVETSGICVYKAAVDPRVTRWDNLQAGVASESVTKIATLQOOSV
OLVKTNNNAVLFKGAKLVDGLSLGLIARNMSSGVYAPILRLAQMODFOOVL
SVRIGELINTTELPGSKMPLPIQGRITFERLLIFRPPDPVLAGIDLDIOCEV
IGIVRSGSGKSTLKLQVMYPERGRVLDSDPLADPAMLRQIGVYQENPLF
NKSRENAMADPGIDLEIVHAATLAGHPRTALPEGYDKICEHGTGLSGGGROR
IAIARALIGDPRIIFEDENTSADYESHPAVNMNRALCKGRTVLIIAHRISTVRAD
RIVVIDKRIYSSGSHETLLNNNGHYAHLRLQGGTT"
/396.3817
/gene="XF2398"
/396.3817
/note="similar to SP|P06739 (percent identity: 39 %/query
alignment coverage: 101.1 %/subject alignment coverage:
100.0 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hemolysin secretion protein D"
/protein_id="AA085197.1"
/db_xref="GI:9107581"
/translation="MKHMLQGRDLSRLRYLAFPAANSIRAOLEPPHRTADETLPA
HLELTPVSPPTARMSMRITLACITVALINACRGTETIVAVAPKIVDSTKYLQPL
ETAVVRILIRDGKRYKAHALLIEDVATINAEYTOADNLNPKTILRLALSTAL
ERQOTPOLTSATIPRADIAEALTLQSNAPQAKHNLQATIAORRAELRTQDPT
LSPAESRIRKISADYAILLKQYVGRHYLLREORITAEEDLTAORNRIOGIRSA
LSAQAELSVLQADRFQOTLDLRKADQIGLTPOLAKKORNRIMTLSPVNGIYO
QIAIHVGVVTPAQQLLAVVPOETLEEVATVYNKQIGLFRQRTVATIESPIYR
YGLTGKVVSISSHDAODEKGLVFPRIRLMDNTLTITDGTQIRMSAGMLSAEIKTG
KRTVIDYLSPLQYIDEALER"
/387.4112
/gene="XF2399"
/387.4112
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA085198.1"
/db_xref="GI:9107582"
/translation="MPINTTASHKTHRPQCOHTCDASNTATPTPTPTPSPL
KDSRASPILAHLCIHPPTKQIHNAHVPPYITK"
/4428.5057
/gene="XF2400"
/4428.5057
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA085199.1"
/db_xref="GI:9107583"
/translation="MTMNTAVFKILKGLQMIYALIMAFGAFISYGIYIKKAAQ
DAIAOKIKESNEILMLOSDSSYKRCFEDELYATGLHGIIFYFNNPLKISYEVK
SPSKPHGPNSSVYIGFGKIVYKKNKINYNKDONOONQYVIGQGLLIERFSNNT
YSAMDKNDIFIEYKADGQPKNKRIHEKLIIIDLEAHTIPQGS"

```

---

```

gene
5054.5326
/gene="XF2401"
5054.5326
/gene="XF2401"
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA085200.1"
/db_xref="GI:9107584"
/translation="MTMSSYFYFDOATLDRAEVYRSSGCPGYQLISAVKAGGDA
RLANGKMGDHSAGRGDPTPSIRAMADQADRENNKINTQHPAL"
/5329.5946
/gene="XF2402"
/5329.5946
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA085201.1"
/db_xref="GI:9107585"
/translation="MNSPRVFKSLKGLALTYLAWVAGATISGKNNMETTTQRA
AAODIKESKMMMLQSRNSSDKRCRFPPEVPIOPSKATFRNIGLISGISTSP
GNLGEAGTSIVALLFPDPKSKKEYITHRLPAGSMQOVIMKEYTFYFYNLDEFLYS
AIDTKGLNIEYQPOQPRNKQEMTHDLOLIDFVMTHTIPQGS"
/5943.6215
/gene="XF2403"
/5943.6215
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA085202.1"
/db_xref="GI:9107586"
/translation="MTMSLTFPOATLERANTYNNNGMAMAYQTLSDAVNAGGDA
RLANGFRNAGIHSPORRDPISIAAMDQADRENNKINTQHPAL"
/6218.6841
/gene="XF2404"
/6218.6841
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA085203.1"
/db_xref="GI:9107587"
/translation="MNSPRVFKSLKGLALTYLAWVAGATISGKNNMETTTQRA
AAODIKESKMMMLQSDHSYKRCRFPPEVPIOPSKATFRNIGLISGISTSP
GNLGEAGTSIVALLFPDPKSKKEYITHRLPAGSMQOVIMKEYTFYFYNLDEFLYS
AIDTKGLNIEYQPOQPRNKQEMTHDLOLIDFVMTHTIPQGS"
/6838.7110
/gene="XF2405"
/6838.7110
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA085204.1"
/db_xref="GI:9107588"
/translation="MNSPRVFKSLKGLALTYLAWVAGATISGKNNMETTTQRA
AAODIKESKMMMLQSDHSYKRCRFPPEVPIOPSKATFRNIGLISGISTSP
GNLGEAGTSIVALLFPDPKSKKEYITHRLPAGSMQOVIMKEYTFYFYNLDEFLYS
AIDTKGLNIEYQPOQPRNKQEMTHDLOLIDFVMTHTIPQGS"

```

---

```

Query Match
Best local similarity 17.3%; Score 86.8; DB 1; Length 15552;
Matches 163; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
QY 62 acccgagagagctgctctcctcgccgagcgcgaatgtagagagccgcatct 121
DB 2498 ACCGCCGATGAGCGACGCTTTCACCCGCACCTAGAACTGACGAAACCCCTGCTCC 2557
QY 122 cccacgcgagacacccgcgcgcctgctgctgcttctacgcgcgtgacgtg 181
DB 2558 CCTACCGCACGCGTGGAGCATGGGATCATTTATGCTGCTCATGCTGACCCCTACTGTGG 2617

```

[illegible][illegible]



```

sequence"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/db_xref="GI:6900416"
/translation="MTLLNLMIMQDYGISVCLTLPYLOHELFESANKSYFSKYLIPVS
LEFTPLSPISVSAFTLPKAMRAAOHSDFOASHYQDRAVARAOOAKAALPLHVS
NASYKOPSPISSTRETPGMSVQVQGLDPAKFAQYROSREPTQAEORFADREEL
LKAASSTYFNVLSPRTVAHAHAKKAAQAOVQAOLFNKGAATALDIEHAKAGYDN
ALAOETALVALEKOTYENQNDLTGIDSKOEAIDETNLAIRYLPKERSLDEWORA
LSNNHEYRMQOALQSSGALRAQNSRTFTYSAHYGYONMTSSGANNNDYRKG
MSYVOLNPLRYTGLGELSKTIEHAQYGAELQALATERHITLALVQATTEGSAAY
OIAQERVLSSRLKLTSTETGQOYGIHRLNLEVIKRAQEVAAQEKLAQARIFMLAY
LRLVKSGLGLETVEFAE"
complement(5707..6042)
/gene="rse1"
/complement(5707..6042)
/gene="rse1"
/complement(5707..6042)
/translation="homology to putative transposase of CAB44496;
designated NM1992 in the Sanger Center 22491 genome
sequence"
/codon_start=1
/transl_table=11
/product="putative transposase"
/protein_id="CAB72028.1"
/db_xref="GI:6900417"
/translation="MSAOVGNRLIAPMYONTMIGVEFAMFOOCLLPALTQKSVILL
DNARFHMVGLREMAEKLYKVLPLAPLPLNIEKWMANIKRYLRTVLSDVARFDD
ALLSTFDPN"
6514..6619
/translation="homology to putative transposase"
/complement(6787..7824)
/gene="relA"
/complement(6787..7824)
/gene="relA"
/function="synthesis of ppGpp, a mediator of the stringent
response"
/translation="homology to GMP-pyrophosphokinase of P1185;
designated relA (NM1991) in the Sanger Center 22491 genome
sequence"
/codon_start=1
/transl_table=11
/product="GMP-pyrophosphokinase"
/protein_id="CAB72029.1"
/db_xref="GI:6900418"
/translation="MAESCKEDLAAPFKTELFPNDIIVLTPHGXVLSLPTGATPIDFA
YALHSIGRCRGARVGEQIVLPSTPEGQVREITITAKEGHSVNMVIEGVKSNKA
IGIRAYIROQNAADVREGVQVLDKQLKLPKPLQELANLGGKKREDLXTAVQ
GEISNRALQKACGLTNEPPVSEITIVKQSKIRKGGNGVLDGSDLMTTLACG
KRAPDDIVGFTVRDGISSVHRKTCPSFRLHAEKVLADSWMALDGGVFAVDIE
IKAOBSGLRVDALNRHKNLNTAVQOSDLEKSMRFTLEVQVDTLPRVLSLIG
DVKGLVSTRL"
BASE COUNT 1853 a 1854 c 2065 g 2052 t
ORIGIN

Query Match 17.2%; Score 86; DB 1; Length 7824;
Best Local Similarity 54.8%; Pred. No. 1.4e-07;
Matches 170; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

OY 62 acccgagacagctgctctcctcgcgccgcgcgaatgtcgagagcgccatct 121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2606 ACGGCGGAAGAACGAGGTTTTCGCCGCAATTGGAACTGACCATACCCGCTCT 2605

OY 122 cccacgcgagactcaagcgccctgtgctgtcttactactcgccgctgagctg 181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2666 GCCCCTCCGAATAGGCGGCGGCTTTATATATGCGGTTTGGCTTTGGCTTTGTGCG 2765

OY 182 gcgggctcgcagatcgatcgttcttctcgcacagaaagatcgctgcggcgac 241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2726 TCCGCTTCGGCAAAATGATATGTGTGCGCGCTTCGCAAAACGATGTCGCGCGG 2785

sequence"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/db_xref="GI:6900416"
/translation="MTLLNLMIMQDYGISVCLTLPYLOHELFESANKSYFSKYLIPVS
LEFTPLSPISVSAFTLPKAMRAAOHSDFOASHYQDRAVARAOOAKAALPLHVS
NASYKOPSPISSTRETPGMSVQVQGLDPAKFAQYROSREPTQAEORFADREEL
LKAASSTYFNVLSPRTVAHAHAKKAAQAOVQAOLFNKGAATALDIEHAKAGYDN
ALAOETALVALEKOTYENQNDLTGIDSKOEAIDETNLAIRYLPKERSLDEWORA
LSNNHEYRMQOALQSSGALRAQNSRTFTYSAHYGYONMTSSGANNNDYRKG
MSYVOLNPLRYTGLGELSKTIEHAQYGAELQALATERHITLALVQATTEGSAAY
OIAQERVLSSRLKLTSTETGQOYGIHRLNLEVIKRAQEVAAQEKLAQARIFMLAY
LRLVKSGLGLETVEFAE"
complement(5707..6042)
/gene="rse1"
/complement(5707..6042)
/gene="rse1"
/complement(5707..6042)
/translation="homology to putative transposase of CAB44496;
designated NM1992 in the Sanger Center 22491 genome
sequence"
/codon_start=1
/transl_table=11
/product="putative transposase"
/protein_id="CAB72028.1"
/db_xref="GI:6900417"
/translation="MSAOVGNRLIAPMYONTMIGVEFAMFOOCLLPALTQKSVILL
DNARFHMVGLREMAEKLYKVLPLAPLPLNIEKWMANIKRYLRTVLSDVARFDD
ALLSTFDPN"
6514..6619
/translation="homology to putative transposase"
/complement(6787..7824)
/gene="relA"
/complement(6787..7824)
/gene="relA"
/function="synthesis of ppGpp, a mediator of the stringent
response"
/translation="homology to GMP-pyrophosphokinase of P1185;
designated relA (NM1991) in the Sanger Center 22491 genome
sequence"
/codon_start=1
/transl_table=11
/product="GMP-pyrophosphokinase"
/protein_id="CAB72029.1"
/db_xref="GI:6900418"
/translation="MAESCKEDLAAPFKTELFPNDIIVLTPHGXVLSLPTGATPIDFA
YALHSIGRCRGARVGEQIVLPSTPEGQVREITITAKEGHSVNMVIEGVKSNKA
IGIRAYIROQNAADVREGVQVLDKQLKLPKPLQELANLGGKKREDLXTAVQ
GEISNRALQKACGLTNEPPVSEITIVKQSKIRKGGNGVLDGSDLMTTLACG
KRAPDDIVGFTVRDGISSVHRKTCPSFRLHAEKVLADSWMALDGGVFAVDIE
IKAOBSGLRVDALNRHKNLNTAVQOSDLEKSMRFTLEVQVDTLPRVLSLIG
DVKGLVSTRL"
BASE COUNT 1853 a 1854 c 2065 g 2052 t
ORIGIN

Query Match 17.2%; Score 86; DB 1; Length 7824;
Best Local Similarity 54.8%; Pred. No. 1.4e-07;
Matches 170; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

OY 62 acccgagacagctgctctcctcgcgccgcgcgaatgtcgagagcgccatct 121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2606 ACGGCGGAAGAACGAGGTTTTCGCCGCAATTGGAACTGACCATACCCGCTCT 2605

OY 122 cccacgcgagactcaagcgccctgtgctgtcttactactcgccgctgagctg 181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2666 GCCCCTCCGAATAGGCGGCGGCTTTATATATGCGGTTTGGCTTTGGCTTTGTGCG 2765

OY 182 gcgggctcgcagatcgatcgttcttctcgcacagaaagatcgctgcggcgac 241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2726 TCCGCTTCGGCAAAATGATATGTGTGCGCGCTTCGCAAAACGATGTCGCGCGG 2785

OY 242 cgtgtaagctggttcagccgcctcgagagctggtgcgtgtgctgcggccatcgtccgcat 301
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2786 CGGAGCAAAACCATCCAGCCGCTGGAACAGTGTGTGTTAAGCGGATGTCGCGAC 2845

OY 302 ggcacaacgcgtcaagcgccgcgagatctgagatcgagctgatacctcgcgggtgtgtgt 361
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2846 GGGCAGCATGTGGAACAGGAGAACCGTGGCGGAACTGGAGGCTGTGGAAACAGACAGC 2905

OY 362 gatgtgcgc 371
||| ||| |||
Db 2906 GATGTGTCGT 2915

RESULT 8
AE002524/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis MC58.
Neisseria meningitidis MC58
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1 (bases 1 to 11381)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uitterback, T.R., Khouiri, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Mastignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
107103075
2 (bases 1 to 11381)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uitterback, T.R., Khouiri, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Mastignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
1. 11381
Location/Qualifiers
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/feature="serogroup: B"
165..976
/gene="NMB1736"
/translation="This region contains an authentic point mutation,
causing a premature stop, and is not the result of a
sequencing artifact; This region contains an authentic
frame shift and is not the result of a sequencing
artifact; similar to PID:1653908 PID:1651692"
complement(1140..2543)
/gene="NMB1737"
/complement(1140..2543)
/translation="This region contains an authentic point mutation,
causing a premature stop, and is not the result of a
sequencing artifact; This region contains an authentic
frame shift and is not the result of a sequencing
artifact; similar to PID:1139569 percent identity: 48.26;
identified by sequence similarity; putative"
/codon_start=1

```

```

/translation=1
/product="secretion protein, putative"
/protein_id="AAFA2081.1"
/db_xref="GI:7226992"
/translation="MTLLNIMMODYISVCLTLPYLQHELFSAKSYSKYILPVS
LFTPLPLSPVSAFTLPEAMRAAOQSHADFOASHYORDAVRARQOAKAFLHVA
NASYORPSPISSTRETQSVQVQGLFLPAKPAQYROSQEDTQAORFADAREL
LKVAESYFVNLISRDYVAHAHAEKAYAOVROALFNKGATALDIHKAQYEL
ALAOEIVALEKOTYENOLDYDLDSCGALIDTALILARYLPLKFRYSIDEMORIA
LSNNHEYRMOQIALOSSGALRAAONSRYPSVSHVYQNNLTSSQNDYHRYGK
MSYGVOLNPLTYGELSGKIHAEAYGAAEOLITERRIKLAVROAYTESGAARY
QIMAOERVLSSRLKSTEGQYQYGRNLVLEIRAOEVAQAEOKLAORXKEMLAY
LRLVESGLTEVFAE"
complement(2612..4039)
/gene="NMB1738"
/gene="NMB1738"
complement(2612..4039)
/gene="NMB1738"
/translation="similar to GP:558154 percent identity: 65.47;
identified by sequence similarity; putative"
/codon_start=1
/translation=1
/product="secretion protein, putative"
/protein_id="AAFA2082.1"
/db_xref="GI:7226993"
/translation="WFSALKSFLSRYIIVRWVAVRDLKPKRTAEBOAFPLAHL
ELDTFVSAPKWAARIMAFALLALIMSWFGKIDIVAASGTVSGRSKTIQPLET
AVKAVHVRDQGHQKQETLAELEAVGSDSDVQSEALQAOLSKRYEAVLALES
RIVPAHIDMAQARSLGSDADVOAQAQVLAHQYMAAOQALOSALRGHAEIOSAKA
OEOKLSVGAIEEQKTADYRRLRADNFIHAFLEOOSKSVNMNDLESFRCQMRQIO
AATAOAEONRVLTQNLKRDITDALRQANQIDQIRQDIDKAKOQOQLTQSPADGT
VOELATYTGCVQAQAKMVAIPDDKMDVEVLNKGDFVEQODAVKLESPY
TRIGYLTGVKVSVDVAQSHQGLVYTAVSLDKHTLNIDKAVNLTAGMNTAETK
TGRRLVDLTLSTQTKLDESFRER"
4275..4628
/gene="NMB1739"
4275..4628
/gene="NMB1739"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation=1
/product="hypothetical protein"
/protein_id="AAFA2083.1"
/db_xref="GI:7226994"
/translation="MTINDLHFLFLGALILPSIGYGTMLLMLVFAVAKPKRN
PKLISGLTAAAILMLFVWILPKQFGPIKEIQAEEDRKYEAEAVFNEQCKTAG
ERTTRRTWMLKIC"
4752..5162
/gene="NMB1740"
4752..5162
/gene="NMB1740"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation=1
/product="hypothetical protein"
/protein_id="AAFA2084.1"
/db_xref="GI:7226995"
/translation="MGNTAMWMSCNPTIPIELFGIOVKIIPLNQIFNHIIPARYTPE
NNVDSKLRRHWAGATIRIDRODEVIAKTIYFEEKLDTGARGMRWKAIIICNK
ERTISEPLSDVLSVLKPYILRPIYASLRDND"
5233..6575
/gene="NMB1741"
/translation="this region contains an authentic frame shift and
is not the result of a sequencing artifact; conserved
hypothetical protein, authentic frameshift; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
6620..6910
/gene="NMB1742"
6620..6910
/gene="NMB1742"
/translation="hypothetical protein; identified by Glimmer2;
putative"

```

```

putative"
/codon_start=1
/translation=1
/product="hypothetical protein"
/protein_id="AAFA2085.1"
/db_xref="GI:7226996"
/translation="MNIOLQGHIVGKKINQIEGSPDYCLIVAPLDSGQNALG
SSTTEYDGSANFQEFNMAQPIEALNVEIVTGTQKLKIVIGFLYKRG"
6915..7112
/gene="NMB1743"
6915..7112
/gene="NMB1743"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation=1
/product="hypothetical protein"
/protein_id="AAFA2086.1"
/db_xref="GI:7226997"
/translation="MQRVYVQSVSTGDFLYLSPETGDIHTKLITNADYFPEEAI
NAGLEETGNQYEFVFGFLMD"
7120..7404
/gene="NMB1744"
7120..7404
/gene="NMB1744"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation=1
/product="hypothetical protein"
/protein_id="AAFA2087.1"
/db_xref="GI:7226998"
/translation="WFGRLKNAPSIITAKHLENKMFINTCRKYGAKLAVVTAAPL
ALAAHANATLPTDPAKNALAKADGMEAGWIVGIPALFVSVKVMK"
7411..7689
/gene="NMB1745"
7411..7689
/gene="NMB1745"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation=1
/product="hypothetical protein"
/protein_id="AAFA2088.1"
/db_xref="GI:7226999"
/translation="MYQVQNCLEKHOENLYSLVVPRIKNGOIVPEYNGSLMK
MSDQPLRLLECSPEKMDIOSGLFTGWIVFGILASVYVSLKRYLK"
7817..8107
/gene="NMB1746"
7817..8107
/gene="NMB1746"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation=1
/product="hypothetical protein"
/protein_id="AAFA2089.1"
/db_xref="GI:7227000"
/translation="MNKNERDEFYISNSDLKLSSEYDPRLPYVYCYLAKETGLKN
FSMDCHNFENINENSCFEITFKDPSFIIGNKIDVSDSNFESVSFEC"
8148..9653
/gene="NMB1747"
8148..9653
/gene="NMB1747"
/translation="similar to GP:3758871 percent identity: 86.60;
identified by sequence similarity; putative"
/codon_start=1
/translation=1
/product="tsap protein, putative"

```

Query Match 17.28; Score 86; DB 1; Length 11381;  
 Best Local Similarity 54.88; Pred. No. 1.3e-07;  
 Matches 170; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

[illegible][illegible]

```

repeat_unit
score: 78.71"
/label=Correia
562..604
/notes="Correia element; hmms hit to HMM Correia (114 -
156), score: 51.92"
/misc_feature
/label=Correia
complement(667..676)
/notes="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
complement(668..1813)
gene
/genes="dapE"
complement(668..1813)
/genes="dapE"
/EC.number="3.5.1.18"
/notes="NMA1730, dapE, succinyl-diaminopimelate
desuccinylase, len: 381aa; similar to many eg. SW:P24176
(DAPE_ECOLI) dapE, succinyl-diaminopimelate desuccinylase
from Escherichia coli (375 aa) fasta scores; E(): 0, 56.9%
identity in 371 aa overlap. Contains Pfam match to entry
PF01546 Peptidase_M20."
/codon_start=1
/transl_table=1
/product="succinyl-diaminopimelate desuccinylase"
/protein_id="CAB8498.1"
/db_xref="GI:7380372"
/db_xref="SPTREMBL:Q9JTL0"
/transl_table="METQSLSELAKALISRPSYTPDDRCQKLVLYITGFAEEL
HGDFENILRGTGVVPCFAGHTDVVPGVVKMDSPEPTERRGRIGRGADM
KTSICEVACICERFAEHPHDSIALITIDDEGDALDGTWKVDYLRKGLIDVC
TYGPEAVADKIDMTKNGRGRSLSGNTLVKKGCHIYPHLAINPTTFAPALLEIQT
EVMDEKNGPEPTSPQISNNGTATNTVITGELNVKPNPEPTSEAGKQKRVHAI
LNHGQYIDLOMSCSGQPEPLQAGKTLQAKTVARAAIAETGCIIEASTSGNSDGFKA
IAKELIEGPNATTHQINENVRDLDPFKLNAVYEGILARLLAGNAV"
/misc_feature
complement(836..1786)
/genes="dapE"
/notes="Pfam match to entry PF01546 Peptidase_M20, score
252.50, E-value 5.7e-72"
complement(1863..1872)
/misc_feature
/notes="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
complement(1927..2547)
/genes="NMA1731"
complement(1927..2547)
/genes="NMA1731"
/notes="NMA1731, conserved hypothetical protein, len:
206aa; similar to SW:P44013 (Y552_HAETN) hypothetical
protein from Haemophilus influenzae (207 aa) fasta scores;
E(): 0, 59.6% identity in 203 aa overlap."
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="CAB8499.1"
/db_xref="GI:7380373"
/db_xref="SPTREMBL:Q9JQ23"
/transl_table="MLTPKSCDLFNIPFQESQLKQVPESIPQIKADYKNOIMQO
LIQVAAELGAPAEPAHPHERMCNGQVRAHFPAKYQYQYNSAALISILNRRRSV
SLUMCHYKADVSPIALPDYNRWLDNFEDEKYSPFMHMGAESEYDDYRTVAQONGESDR
KLQNDDEFCIGKIHTRDLDGRQDVAKMIAETVEDLLPLYRACHGK"
misc_feature
complement(2551..2558)
complement(2570..2579)
/notes="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
complement(2581..3090)
/genes="NMA1732"
complement(2581..3090)
/genes="NMA1732"
/notes="NMA1732, conserved hypothetical protein, len:
169aa; similar to SW:P43971 (Y243_HAETN) hypothetical
protein from Haemophilus influenzae (172 aa) fasta scores
E(): 0, 49.7% identity in 169 aa overlap."
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"

```

/translation="MLTNIARKIFGSRNDRLKQYRKSVARINALEBOMOLASDQIO  
AKTAEFKORLADGOTLIDGILPEAFACGRASRSTLGMRHQDQIMELNYLGLTIVGIIISD  
RTGGKTLVATLAVLYNLNLAGGVHVVYVNDYLASDAGIKMELNYLGLTIVGIIISD  
MOPEDRONAVAADITYGNNNEGFEDILRDMNYTDQDYKORELNFVAVDEVDLILDE  
ARTLILISGQADNLIQLOIYIMNTVPPHILVORTFEEBGEVGYWDEKHOVILTSAGHEH  
AEQILITGQDLAENDISYSAANIALMHINIMALIRATLLEPHDKHUYIODEIYIYVEF  
TGRMSGRNMESEGLHOAVEAKEGVELKRENOLASTIPFONYFLRYLILKSMGTADFE  
AFEEOSIYNLETVILPTNRVYORKDENDIFSAEKKFAVYDEIECHKRQGPVLYG  
TTSIENSELYSHLQKAGLPHNYVLAKHEBERALIVAGQKGYATIVATNMGRTDII

Query Match	Best Local Matches 170;	Similarity	Conservative	Score 86;	DB 1;	Length 326301;	Pred. No. 8.3e-08;	Mismatches 140;	Indels 0;	Gaps 0;
QY 62	accgagacagagctggcctctctccgagccgcctcgaatgtgcagagccgcacatc	121								
Db 273549	ACGGCGGAAAGAACAGCGCGTCTTTTGGCCCGGCAATTGTGAACTGACGATACGCCGCTCT	273490								
QY 122	ccacacgagagactcaccgcccctctgtctgtctgtcttcttaactgcgcgctggcgttg	181								
Db 273489	GCCGCTCCGAAAGGGCGCGCGCTTTTATTAATGCGCTTTTCGCTTTGGCTTTGGTGTGG	273430								
QY 182	gcggagctgcgcagagatcgcacatcgttctgtctgtcatccagaagaatctgcgcggagac	241								
Db 273429	TCCTGGTTGGGCAAAATTCATATTGTGGGGCGCGCTTCGGGCAAAACGTTGCGGCGGG	273370								
QY 242	cgtgtaaagctgtgttaaccgcctcgaagtcgcgcgtgtgtgtgcacactaatgtccgcgat	301								
Db 273369	CGAGCAAAACCATCCACGCCGCTGGAACAGGTGGTGTTAAGCGGTCATCTGCGCGCAC	273310								
QY 302	ggccaacacgtcaagccgcgcagatctgtatcgcagctgtatccatctgcgcggtgtgtg	361								
Db 273309	GGGCAACATGTGAACACGAGGAAACGCTGCGCGAATGTGAAGCTGTGGAAACACACACAC	273250								
QY 362	gatgtttgcgc 371									
Db 273249	GATGTGTGCT 273240									
RESULT 10	AX044033/C									
LOCUS	AX044033	349980 bp	DNA	linear	PAT 24-NOV-2000					
DEFINITION	Sequence 112 from Patent WO0066791.									
ACCESSION	AX044033									
VERSION	AX044033.1	GI:11342917								
KEYWORDS										
SOURCE	Neisseria meningitidis.									
ORGANISM	Neisseria meningitidis.									
REFERENCE	1 (bases 1 to 349980)									
AUTHORS	Piazza,M., Hickey,E., Peterson,J., Rettell,H., Venter,J.C., Masiagiani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M., Scariotto,V., Rappuoli,R., Frerzer,C.M. and Grandi,G.									
TITLE	Neisseria genomic sequences and methods of their use									
JOURNAL	Patent: WO 0066791 A 112 09-NOV-2000;									
FEATURES	CHROM. COORDINATE (US) : THE INSTITUTE FOR GENOMIC RESEARCH (US)									
SOURCE	Location/Qualifiers									
	1..349980									
	/organism="Neisseria meningitidis"									
	/db_xref="taxon:487"									
	/note="sequence too long, cut in 8 pieces.-seq 1: 1 to 349980 349980 bases-seq 108: 300001 to 649980 349980 bases-seq 109: 600001 to 949980 349980 bases-seq 110: 900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980 349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq 113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to 2272325 172325 bases"									
BASE COUNT	87189 a 93501 c 84627 g 84663 t									
ORIGIN										
Query Match	17.2%;	Score 86;	DB 1;	Length 326301;						
	Best Local Matches 170;	Similarity	Conservative	Score 86;	DB 1;	Length 326301;	Pred. No. 8.3e-08;	Mismatches 140;	Indels 0;	Gaps 0;





predicted by Homology  
 /product="PUTATIVE VGR-RELATED PROTEIN"  
 /protein\_id="CAD17326.1"  
 /db\_xref="GI:17430643"  
 /translation="MDEKTLQHLFAPAHRLYALSGECPILAELAVAWLIGRELSLEF  
 EWRVVASANARFALSEPTGQVTVITLADSTQARFGLIRQAQLDAGSLARL  
 TYPWPMFATQORHQVQONRPLADILQVLSVAPYAMRPAAGADMDAGSLARL  
 IAQRETRYREYTRLLAEGISFTTVEDEOAPSGHALLIFADSPRLAEDTASAGGI  
 RTHRAHQSEDAIQAALCHSRTAVGVAVAAADAKARAFARHPRRCGIAAGSDA  
 YLSISPLAPDANARQIAEOVMESVBARFLFIRGSVTRTSGSTRKLYVDCPLPK  
 DVGGPYPLDLVEHCGINNLTADTQATLARLGGIDALTTDTTPSPSPESGPGFG  
 MAPHEVIERETPTADLTAAARAHGAOPFRAGDARPRMVPVSHPTGRLYAETPARG  
 VOSATVVGPNGETEASGDAEHHTTPRGQYRFRFQWQGERPDRSTGWLVAQRAGA  
 GMQWMLPRIGOEVIYKFOEDTIDQVIGALYNGQEGAGIAPTPGQVAGVREGAD  
 LVFDSDQLRTQATQAYSQLNGLIHQODNRGSEFRGQFELRTDYGAVRGSA  
 GVLITRYDATSGKAVPTGDNAAGIALIROAKOLTAISGQATTHOTATLSTANDSA  
 PLAOEKASGMVQKALDAQOADAAGNTQCKVPHQSEATVQIARAGLAVIAGO  
 DLOFANGESVYLASGQDTNVAVGKQARHAGQIGVAGLSKAGDNDIGLDTAGQDN  
 LDVOADALGILSKDILKIVASNLHDPRAAKRIRLATAAGASITLEGGINIVETPG  
 RITYKAORSPAGVYQSYPLPVPQGVCECVLRAMGKHAFANKNG"  
 1501. 2865

repeat\_region  
 /note="RS50+ or RS06148  
 PROBABLE LONG DIRECT REPEAT 50 +  
 predicted by Homology"  
 /evidence=not\_experimental  
 2999. 3835  
 /gene="RSP0176"  
 /note="RS04688"  
 2999. 3835

gene  
 /function="miscellaneous; hypothetical/global homology"  
 /note="Product confidence : hypothetical  
 Gene name confidence : hypothetical  
 predicted by Codon\_usage  
 predicted by Framed"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="CONSERVED HYPOTHETICAL PROTEIN"  
 /protein\_id="CAD17327.1"  
 /db\_xref="GI:17430644"

gene  
 4412. 6634  
 /gene="RSP0177"  
 /note="RS04689"  
 4412. 6634  
 /gene="RSP0177"  
 /function="miscellaneous; hypothetical/partial homology"  
 /note="Product confidence : hypothetical  
 Gene name confidence : hypothetical  
 predicted by Codon\_usage  
 predicted by Homology  
 predicted by Framed"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="HYPOTHETICAL PROTEIN"  
 /protein\_id="CAD17328.1"  
 /db\_xref="GI:17430645"

gene  
 /translation="MSYLELGPNPALTVDEQARFLACRMEQEFHCSTEFHLSFF  
 DGTNNKTRDTPROASHNVARLFDIPEEOHRIIYVGIPTPEKETGSDGRDAR  
 AGTAGGGEARIMWALLOVNTALYAYYTQSLSMAGVDELSTVHOTSSLNMGICA  
 VTAGRDEVEQLANAKNSEMYITAAETVLPFRHBERALLRQREYISLTKALING

gene  
 6745. 7527  
 /gene="RSP0178"  
 /note="RS04690"  
 6745. 7527  
 /gene="RSP0178"  
 /function="miscellaneous; unknown"  
 /note="Product confidence : hypothetical  
 Gene name confidence : hypothetical  
 predicted by Codon\_usage  
 predicted by Framed"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="HYPOTHETICAL PROTEIN"  
 /protein\_id="CAD17329.1"  
 /db\_xref="GI:17430646"

gene  
 7518. 8021  
 /gene="RSP0179"  
 /note="RS04691"  
 7518. 8021  
 /gene="RSP0179"  
 /function="miscellaneous; hypothetical/partial homology"  
 /note="Product confidence : hypothetical  
 Gene name confidence : hypothetical  
 predicted by Codon\_usage  
 predicted by Framed"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="HYPOTHETICAL PROTEIN"  
 /protein\_id="CAD17330.1"  
 /db\_xref="GI:17430647"

gene  
 8488. 8488  
 /gene="RSP0180"  
 /note="RS04692"  
 8488. 8488  
 /gene="RSP0180"  
 /function="miscellaneous; hypothetical/global homology"  
 /note="Product confidence : hypothetical  
 Gene name confidence : hypothetical  
 predicted by Codon\_usage  
 predicted by Homology  
 predicted by Framed"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="CONSERVED HYPOTHETICAL PROTEIN"  
 /protein\_id="CAD17331.1"  
 /db\_xref="GI:17430648"

gene  
 8956. 8956  
 /gene="RSP0181"  
 /note="RS04693"  
 8956. 8956  
 /gene="RSP0181"  
 /function="miscellaneous; hypothetical/partial homology"  
 /note="Product confidence : hypothetical  
 Gene name confidence : hypothetical  
 predicted by Codon\_usage  
 predicted by Homology  
 predicted by Framed"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="HYPOTHETICAL PROTEIN"  
 /protein\_id="CAD17332.1"  
 /db\_xref="GI:17430649"

RPKRMICIRLTVSGRGAARVPRANMLKDALDDMTLAGVSPDFLIGDPTVASV  
 GVANSKTVATGHSGWCEAEFLRIPGVYKRVYLVSAHEVSGSPFLDKAAVDNCLLAY  
 PGVHTDVGAAYQPGDGRGCTGDDKPDNSKLSQITLAKMYRPAAGVPLNDMANLT  
 TPEVRALKISPLIOAYNDYVAVNPLIRKHGGGTGAAYQGYLYLMRRMRLAG  
 AAEFENOLFQRAEFOYGACANDLVANLLIREAKDLAARENDPAAYAGMQRORV  
 LPVTOGVAWNGVKQKQWGDQKVRERREYAVNNDRPLDPRYTRICDDYTHSRAMEK  
 PFGAPSDSVMRBQOARLEOLKQOAAKQVADLSRDLIGTLKRYESTYANGDEEA  
 PNVYVQDRDLIOYKDKDSVYMETRGREPSMWGLRMRTHPAPPTLGERAOAMD  
 EVASVGRKAKAQAQEAHDAEARKLANTARKLLEAGQESLERATDALQRFLLGGVRF  
 6745. 7527  
 /gene="RSP0178"  
 /note="RS04690"  
 6745. 7527  
 /gene="RSP0178"  
 /function="miscellaneous; unknown"  
 /note="Product confidence : hypothetical  
 Gene name confidence : hypothetical  
 predicted by Codon\_usage  
 predicted by Framed"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=not\_experimental  
 /product="HYPOTHETICAL PROTEIN"  
 /protein\_id="CAD17329.1"  
 /db\_xref="GI:17430646"



FASTA scores: E(1): 0, 60.8% id in 227 aa, and to Burkholderia pseudomallei transcriptional activator protein TrIR SW:TrIR_EURS (031395) (229 aa) fasta scores E(1): 0, 57.6% id in 224 aa"									
/codon_start=1									
/transl_table=11									
/product="transcriptional activator"									
/protein_id="CAC90814.1"									
/db_xref="GI:15980015"									
Query Match	13.9%	Score 69.8;	DB 1;	Length 334050;					
Best Local Similarity	51.8%;	Pred. No. 0.00013;							
Matches 158;	Conservative 0;	Mismatches 147;	Indels 0;	Gaps 0					
OY 62	accggagacgagctgcttcctccgcccggccctcgaaatgtcgaagccgcacatc 121								
Db 260521	ACCCGATGATGAATATATATTTTGGCTGTATTATTGGAAATGTTCAGCGCCCGATTTGCC 260580								
OY 122	cccaccgcgagactcaacgcgcgcctctgctctgtctgtctgtcttctcactgcgcggtgagtg 181								
Db 260581	CCATTAGCCCGCGTACCGCGTGATTCACATCACTACATTAATGTGTTGGTATTGATATGG 260640								
OY 182	gggggtctgcggaagatcgacatcgcttgccttcgtcatccagaagaatcgtgcggcgagc 241								
Db 260641	GGGATTATTGGAAACCTTATATCCATGCTTCAGCCAGCGCGCAAGTGATGTGGCGAA 260700								
OY 242	cgtgtaaacgtgtctcagcgcgctcgagtcgagtcgcggtggtgcggccaactatgtccgagt 301								
Db 260701	CATTCCAAATATATCCAACCCGACAGCTGAGTGTGATGACGAAATTAATATCTCCAGAT 260760								
OY 302	ggccaaaccgtcaagccgcgcgagattctgactgcagctgtgatacttcgcgcggtgtgtg 361								
Db 260761	GGGGATACCGCTTGATGAGCGCGCAGGTGCTTATTCATTAAACCCATTGGCATGTATGCC 260820								
OY 362	gatgt 366								
Db 260821	GAAGT 260825								
RESULT 14									
LOCUS	BPCYADE	6645 bp	DNA	linear	BCT 08-SEP-1992				
DEFINITION	B.pertussis genes cyaA, cyaB, cyaD and cyaE.								
ACCESSION	X14199								
VERSION	X14199.1	GI:39731							
KEYWORDS	adenylate cyclase; cya gene; cyaA gene; cyaB gene; cyaD gene; cyaE gene; hemolysin; secreted protein; toxin.								
SOURCE	Bordetella pertussis.								
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae; Bordetella.								
REFERENCE	1. (bases 1 to 2040)								
AUTHORS	Glaser,P., Sakamoto,H., Bellalou,J., Ullmann,A. and Danchin,A.								
TITLE	Secretion of cycloolysin, the calmodulin-sensitive adenylate								
JOURNAL	cyaB-haemolysin bifunctional protein of Bordetella pertussis								
REFERENCE	EMBO J 7 (12), 3997-4004 (1988)								
MEDLINE	89091151								
AUTHORS	2. (bases 1 to 6645)								
TITLE	Danchin,A.								
JOURNAL	Submitted (08-SEP-1992) A. danchin, Institut Pasteur, Unite de								
REFERENCE	Regulation de l'Expression, Genetique, 28 rue du Docteur Roux,								
COMMENT	75724 Paris Cedex 15, FRANCE								
FEATURES	see I00345 for upstream cya gene; cya Operon is organized cyaABDE.								
source	cyaB (712 aa) is initiated 78 bp downstream of cyaA stop; cyaB stop								
gene	overlaps with cyaD initiation (440 aa).								
	Location/Qualifiers								
	1. .6645								
	/organism="Bordetella pertussis"								
	/strain="1833."								
	/db_xref="taxon:520"								
	1. .725								
	/gene="cyaA"								

```

CDS
<1..658
/gene="cyaa"
/codon_start=2
/transl_table=11
/protein_id="CAA32411.1"
/db_xref="GI:39732"
/db_xref="SWISS-PROT:P15318"
/translation="PGRGLDAGAKGVFLSLGKGFASLMDPEFSTSVLLENINAVSAR
DDVLIGDAGAVNLGAGDDVLLSGAGDDVLLSGDNDLLSGDNDLLSGDNDLLSGDNDLL
LFGYGHDTYESGGHDTIRINAGADOLIFARQNDLIRLIGTDLALVHMYRD
ADHREYTIHANONAVDQAGIEKLVENAAQYDPGAAAAAPPAAVPTLMGLAVNMR"
721..725
/gene="cyaa"
736..2874
/gene="cyab"
736..2874
/gene="cyab"
/codon_start=1
/transl_table=11
/protein_id="CAA32412.1"
/db_xref="GI:39733"
/db_xref="SWISS-PROT:P18770"
/translation="MTSPVACASVPDGLCLVWLARYHGLAADPEQLRHEFAEQAF
CSERTQLAARVGLKVRHRRPAPAPALPALADROGGYFVLPPEEGADQAVL
IQRGQAPARLUGQAEFEALMAGELLCCASPTQALARPDSWFTIPALVHRLIGE
VLLISLVQLFSLTPLFFQYVMDVAVNNAAETLVITVGFLLALILDEALTGIRTY
VLAHTSKLDELGLARLYAHLRLPLAVQARVDSVARELHIAFLGNATV
LLDVFVSVIDVAFVSVKLLVLAALPCYFLLSLVTLVLRRLHVKFERGAE
AFIVETVSIDIVKSLAVEPQWRMDROLAGVAAGLSVANAMANTGVLLISRL
PWSGMAHGRGARMTVGLVAFNMLSGHYTOPYIRLAQNDFOQTGVSQMRGDI
INCRTEVYAGRAQDPAIRGSTIELDRSPRYRPDADALRNVSRLTAPLOESTLFPNSVSDI
SGKSTLRILQRMFVADNRVILIDGHDIYDSASLRQGLVIOESLTFPNSVSDI
ALTRGASMEHVAARLGAHEFTCOLPEGDTMLGENGVLSSGQRORICIAALI
HRRPVLIDFATSDADYSEHIIQNRMDICGRVYIIIAHRLSAVRCADRLVMEGG
EVAECGSHETLLAAGLYARLQALQAGEAG"
2862..2866
/gene="cyad"
2862..2866
/gene="cyad"
2871..4193
/gene="cyad"
2871..4193
/gene="cyad"
/codon_start=1
/transl_table=11
/protein_id="CAA32413.1"
/db_xref="GI:39734"
/db_xref="SWISS-PROT:P11091"
/translation="MRRLALRELAARHGVVAASWROHRRRPAQWEDPVTEFLPSALS
LOERPISTPARMLARILMALAGALVWSVKTETIVHAAKVPVPGQSKIIAASLGTG
RYARVIVADNSRYRAGDVLIDAGVTEAEERKWRFOAQAORDEARSAMIATLGTG
RAPVLAELPADPQGMMAOSYILDSOVADVOALRSTAIATYRRDGLVTOIHAHAG
LRNDGVSOOATLEKQAKMTLEGRLOSEAORAAICQITRRQAPEFTILARLAAOA
EOEIAITSAQSRSLVLTAPVDVQOQVLTBETVAATOPTLMVYPSGAGIOVOAOI
DSKDIOFVRAGAPATVKGAVTYTKYGTLEGKVLVYSPDVLVDDROHSTRITALLAH
PALEYDGKRLKEGMVAVODIRFGSRRLLEYLLSPVARRAGESLGER"
4184..4187
/gene="cyae"
4184..4187
/gene="cyae"
4195..5619
/gene="cyae"
4195..5619
/gene="cyae"
/codon_start=1
/transl_table=11
/protein_id="CAA32414.1"
/db_xref="GI:39735"
/db_xref="SWISS-PROT:P11092"
/translation="MAAVYRRGRALALALMAGFALSVGGRARDGLATPPAFEGQ
AAPAVSWPCPPADRLDILALIEALCALCHSPALRGMARIKARSFVGLARVYYP
SVALSGRSRKQRSTGLGDEGVNRNNAVYTLANFLPDSGRSALRAAQAOQLDEAAQA

```

```

misc_feature
5706..>6645
/gene="cyay"
/note="unknown ORF"
5706..6645
/gene="cyay"
5760..6617
/gene="cyax"
5760..6617
/gene="cyax"
/feature="unknown ORF"
/codon_start=1
/transl_table=11
/protein_id="CAA32415.1"
/db_xref="GI:39736"
/db_xref="SPTREMBL:Q45360"
/translation="MRPVHDCGRAAPCVAAVGFOPAPVAGDLGRHASDRPGLPDPD
GAGEOFRSAAAEIYKRYIDARGEGATGHDMSVLPVAMPICLATSRFPAMCARMOVG
GAASVRLHGVNVDHSEVWLAGVVDLVCPRHAYEDTOLDQSVYRACVADVYRLYR
PAGMALDAEALACEAPQARYPRVYADGATYRRLIELSCPLORADQRCVETFPD
GAPHLVRAGVGLAMLPDSSARAAVEHGIYVADRCAIPMETCMYARRHGPQAEAVE
RVFDQFGRC"

BASE COUNT      956 a      2142 c      2441 g      1106 t
ORIGIN

Query Match
Best local similarity 12.5%; Score 62.6; DB 1; Length 6645;
Matches 152; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 65 cgggaagagctgctctctcgcggcgccctcgaatgtcgagacgcgcgcacatccc 124
DB 2967 CCGGTGGAGACCGAGTTCCTCCCTTCGGCTTCCTCGTAGAGAGCTCCGATATGCC 3026
QY 125 accgcgagactcaagcgccgctgctgctgttctactgcgcgctgctgagcg 184
DB 3027 ACGGCCCGTGGCTGGCGCGCATACGATGCGCTGGCGCGCGCGCTGCTGCTG 3086
QY 185 ggtctgagcagatcgacatcgctgtcttgcacccagaagatcgctgagcgagc 244
DB 3087 GTAGTGGCAAGACCGAGTCGTGTCATCGGCGGCGCAAGGTGCTGCGGCGCAG 3146
QY 245 gtaagctggttagccgctcagagctgctgctgctgctgctgctgctgctgctg 304
DB 3147 AGCAAGATCATCGCGGCTCCGAACCGGCTGCGCGCGCTGCGCGCGGACAT 3206
QY 305 caaacgctcaagcgcgagcagatctgactgagctgacatccatcgcggtgtgtgat 364
DB 3207 AGCCGCTGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCGCGCGCGCGGAGCGGAA 3266
QY 365 g 365
DB 3267 G 3267

RESULT 15
AF193064 7493 bp DNA linear BCT 14-NOV-1999
LOCUS AF193064
DEFINITION Caulobacter crescentus truncated S-layer protein (rsaA), ABC
transporter (rsad), and membrane forming unit (rsaf) genes,
complete cds.
ACCESSION AF193064
VERSION AF193064.1 GI:6409374
KEYWORDS Caulobacter crescentus.
SOURCE Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
ORGANISM Caulobacter.

```



Tue Jul 16 05:46:21 2002

us-09-484-577a-3.rge

Page 18



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 22:20:07 ; Search time 1623.58 Seconds  
(without alignments)  
4164.855 Million cell updates/sec

Title: US-09-484-577A-3  
Perfect score: 501  
Sequence: 1 actccagcctctacacga.....cagaatgcgcgatgatcat 501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
1 number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estha:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_ges:\*  
13: em\_ges\_hum:\*  
14: em\_ges\_inv:\*  
15: em\_ges\_pln:\*  
16: em\_ges\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.6	12.1	932	12	CNS0070E
2	55	11.0	925	12	CNS0091P
3	44.8	8.9	508	9	AV630811
4	43.6	8.7	772	12	AF075872
5	42	8.4	939	12	CNS004NB
6	40.8	8.1	932	12	CNS00720
7	40.4	8.1	774	10	BI956591
8	40.2	8.0	581	9	AV921113
9	40.2	8.0	586	9	AV913605
10	40.2	8.0	605	9	AV918908
11	40.2	8.0	659	9	AV931532
12	40.2	8.0	660	9	AV946608
13	39.8	7.9	639	9	AI770858
14	39.6	7.9	574	10	BM349363
15	39.6	7.9	576	10	AI881920
16	39.6	7.9	577	10	BM267007
17	39.6	7.9	645	12	CNS01213

C 18	39.6	7.9	686	10	BI959511	BI959511 HVSMEn001
C 19	39.6	7.9	692	12	CNS007MH	AL050923 Drosophila
C 20	39.6	7.9	754	10	BF859678	BF859678 963014A09
C 21	39.4	7.9	814	10	BG310235	BG310235 HVSMEn001
C 22	39.2	7.8	940	12	CNS0625X	AL421915 T7 end of
C 23	39.2	7.8	870	12	BI950671	BI950671 HVSMEn002
C 24	38.8	7.7	446	10	BM372629	BM372629 EBP105_S0
C 25	38.8	7.7	603	10	BF253745	BF253745 HVSMEn000
C 26	38.8	7.7	660	9	BE215392	BE215392 HV_CED000
C 27	38.6	7.7	844	12	CNS0052P	AL050652 Drosophila
C 28	38.4	7.7	642	10	BI954172	BI954172 HVSMEn001
C 29	38.4	7.7	925	12	CNS0091P	AL050303 Drosophila
C 30	38.4	7.6	265	10	R01662	R01662 yef5a03.s1
C 31	38	7.6	765	10	BF616985	BF616985 HVSMEn001
C 32	37.8	7.5	672	9	BB528893	BB528893 BB528893
C 33	37.5	7.5	424	10	BE404013	BE404013 WHE0410_D
C 34	37.4	7.5	455	9	BB833701	BB833701 BB833701
C 35	37.4	7.5	698	9	BB610391	BB610391 BB610391
C 36	37	7.4	392	9	AV644920	AV644920 AV644920
C 37	37	7.4	393	9	AV644931	AV644931 AV644931
C 38	37	7.4	424	10	BE497481	BE497481 WHE0753_A
C 39	37	7.4	539	9	AV640485	AV640485 AV640485
C 40	37	7.4	629	9	AV912950	AV912950 AV912950
C 41	37	7.4	667	12	A2131633	A2131633 OSJNB011
C 42	36.8	7.3	441	10	BE604637	BE604637 WHE1413-1
C 43	36.8	7.3	510	10	BP585055	BP585055 57F-ZO P
C 44	36.8	7.3	521	9	AV434008	AV434008 AV434008
C 45	36.8	7.3	600	9	AV602736	AV602736 AV602736

## ALIGNMENTS

RESULT 1  
CNS0070E/c  
LOCUS  
DEFINITION Drosophila melanogaster genome survey sequence TETS end of BAC # BAC1421 of RPCI-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.  
ACCESSION AL066254.1 GI:4945121  
VERSION AL066254  
KEYWORDS fruit fly.  
SOURCE Drosophila melanogaster  
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 932)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr

## COMMENT

- Web : www.genoscope.cns.fr  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source  
1..932  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone\_id="RPCI-98"



BASE COUNT	69 a	194 c	172 g	72 t	1 others
ORIGIN					
Query Match		8.9%;	Score 44.6;	DB 9;	Length 508;
Best Local Similarity	52.1%;		Pred. No. 1.2;		
Matches 100;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;	
QY	5	tccgcgcctctcacccgagagatgaatgcgcgcctcgttgaaagtgtgtgcggtctcggggagcaaaccc	64		
Db	331	TCCACACACCGTGAAGGCGGCGCTACGCGCCCGCGCGCGGCGGAGAGCGCGCGCAGCTTTCACG	272		
QY	65	cggagacgagctgtgcgccttccttcgcggcgcgccctcgaaattgtcgagagacgcgcacatctccc	124		
Db	271	AGCGCGCGAGGGGACCCAGCGTGCCGTACACAGTGCGCCGACGAGCCGACATCATGCGAACAACCC	212		
QY	125	accgcgaacactcacgcgcgccttgcttgccttgcttctactgacgcctgagcgtgagggcg	184		
Db	211	TGCGGCGCCCATGAACAGCGCGCCACGTCGCCCGCACCTCGTGCGGCGCGGGAGTGTGGGCG	152		
	185	ggtctctggcagag	196		
	151	GTGATGCCGATG	140		

RESULT	4			
AF075872/c				
LOCUS		772 bp	DNA	linear
DEFINITION	AF075872			GSS 29-AUG-2000
	AF073672, Salmonella typhimurium LT2, Lambda DASH II Salmonella			
	typhimurium genomic clone 14-2T3, DNA sequence.			
ACCESSION	AF075872			
VERSION	AF075872.1	GI:3320742		
KEYWORDS	GSS.			
SOURCE	Salmonella typhimurium.			
ORGANISM	Salmonella typhimurium.			
	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
	Salmonella.			
REFERENCE	1 (bases 1 to 772)			
AUTHORS	Wong R.M.-Y., Wong K.K., Benson N.R. and McClelland M.			
TITLE	Sample sequencing of a Salmonella typhimurium LT2 lambda library:			
	comparison to the Escherichia coli K12 genome			
JOURNAL	PEMS Microbiol. Lett. 173 (2), 411-423 (1999)			
MEDLINE	99243757			
COMMENT	Contact: McClelland M			

FEATURES	source
location/Qualifiers	1. .772
/organism="Salmonella typhimurium"	
/strain="LT2"	
/db.xref="taxon:602"	
/clone="142-T3"	
/clone_1b="Salmonella typhimurium LT2, Lambda DASH II"	
/vector="Lambda DASH II; sequenced using LI-COR sequencer"	
BASE COUNT	169 a 229 c 205 g 167 t 2 others
ORIGIN	

Query Match	8.7%	Score 43.6	DB 12	Length 772
Best Local Similarity	47.4%	Pred. No. 2.6		
Matches 130	Conservative 0	Mismatches 144	Indels 0	Gaps 0
QY	72	agctggcctctctgcgcgcgcctcgaaatgtcgaagcgcgcacatcccaacgcga	131	
Db	410	AGCATGAGCCGCGCGATGACGATCCCGATATTCACGCTGAACGCGCGCTTTCCGCGCGG	351	
QY	132	gaccacagcgccgcttgcgttcgtctactacgcgcgctgagcgaggcgagctccg	191	
Db	350	GTCGATATGTTCTGATCGTCATCGTATTTCTCATTTCTCGGCATCGCGCGTGGTTGG	291	

QY 192 gccgagatcgacaatcgctgtcttcgtgataccagaagaatcgicgcggcgacccgtgtaaacg 251

Db 280 gccgcattcgatgagagattttccacccggcaacgggaagatgattccacgattccacggcgaacgg 231

QY 252 tggattcaacgcgcgtcgaagtcgagtcgctgagtcgagccactcattgcctcgatgtgccaaacg 311

Db 230 tttctcacttcctctgattgcccgcattttctggcgcacattacggtgcgggaagcgcacag 171

QY 312 tcaagcgccgcgagatctcgtatcgaatcgtgatacc 345

Db 170 tttcagcctaaccaagatttctgcgccggccttgcattcc 137

RESULT	5		
CNS004NB/c			
LOCUS		839 bp	DNA linear
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #		
	BACR10E16 of RPCI-98 library from Drosophila melanogaster (fruit		
	fly), genomic survey sequence.		
ACCESSION	AL054280		
VERSION	AL054280.1	GI:4931788	
KEYWORDS	GSS.		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster		

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 839)  
Genoscope,  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

**COMMENT**

Determination of this BAC-end sequence was carried out as part of collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/The\\_BDGP\\_Drosophila\\_melanogaster\\_BAC\\_library](http://www.fruitfly.org/The_BDGP_Drosophila_melanogaster_BAC_library) was prepared by Kazuhiro Oseegawa and Aaron Mamoser in Pletier de Jong's laboratory in the department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPC1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library or filters for hybridization from the BACRAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```

source
1. .839
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_id="RPCI-98"
/clone="BACR10E16"
/note="end : TEL3"
285 a 67 c 77 g 27 t 383 others
BASE COUNT
ORIGIN

```

[illegible]



```

/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="bagsla15"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna

```

BASE COUNT	ORIGIN
134 a	162 c 153 g 137 t



	Query Match	8.0%	Score 40.2;	DB 9;	Length 660;
	Best Local Similarity	45.28;	Pred. No. 16;		
	Matches 147;	Conservative	0;	Mismatches 178;	Indels 0;
				Gaps	0;
QY	88	gagccgacctgaattgtcgagacgcgcacccctccaccgcgagatcaaggccgctt	147		
Db	553	GGAAGCGGAGCTCATACCATTTGAGAACTCACGCTCCGACGAGAACTCCTGTGGTCTT	494		
QY	148	gcttgcgtccttgcttactactgcgcgttgcgttgcgttgcgttgcgcgcagatcagcatcgt	207		
Db	493	CCGCGAAGGGGTGACTACGAGGGCCGGCAATTGGGCCCTACGACATCCACTCC	434		
QY	208	tgccttcgcattccagaaagatcgtgcgcgcgcacacctgtataagcttggttcaagccgctcga	267		

	Query Match	Similarity	Score	DB	Length
Best Local	Similarity 48.18;	Pred. No. 19;			
Matches 113;	Conservative	0;	Mismatches 122;	Indels 0;	Gaps 0;
QY	142	cgccttgctgcgcgccttgcctactgcgcgcgtgcgtgcgcggtctcgcagatcga	201		
Db	567	CGTCGGCCCTTGCAAGTAATCGGACGGTGGGGCTCCGCCGCCGCCCAACGGGGCCCT	508		
QY	202	catcgttgccttcgcgcacagaagatcgtgcgcgcgcgcgcgtgtaagctggttcagcc	261		
Db	507	TGCGGGGGCGTCGTCCTCCGCCGCGCCTTCGTCGCCACAGGGCAGATCTTCAGC	448		
QY	262	gctcgaagtcgcgcgtgcgtgcgcacactatgtccgcgatgcgcgaacacgtcgaagcccg	321		









OM nucleic - nucleic search, using sw model

```
Run on:      July 15, 2002, 23:19:32 ; Search time 205.98 Seconds
              (without alignments)
              4176.008 Million cell updates/sec
```

```
Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctcagcctctcacga.....cagaatgcgcgatgcat 501
```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

```

Searched:      1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

```

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
Listing first 45 summaries

```

Database : N\_Geneseq\_032802:\*

1:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1980.DAT.*
2:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1981.DAT.*
3:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1982.DAT.*
4:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1983.DAT.*
5:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1984.DAT.*
6:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1985.DAT.*
7:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1986.DAT.*
8:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1987.DAT.*
9:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1988.DAT.*
10:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1989.DAT.*
11:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1990.DAT.*
12:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1991.DAT.*
13:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1992.DAT.*
14:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1993.DAT.*
15:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1994.DAT.*
16:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1995.DAT.*
17:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1996.DAT.*
18:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1997.DAT.*
19:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1998.DAT.*
20:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA1999.DAT.*
21:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA2000.DAT.*
22:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA2001.DAT.*
23:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA2002.DAT.*
24:	/SDSI/gcgdata/hold-genseq/genseqn-emb1/NA2003.DAT.*

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	88.8	17.7	28804	17	AAAT37329	Spingina biosynthe
C 2	88.8	17.7	28804	18	AAAT92474	Sphingomonas genus
C 3	88.8	17.7	28804	20	AAAV9812	Sphingomonas 588 s
C 4	88.8	17.7	28804	20	AAV81474	Chromosomal fragme
5	86	17.2	1305	21	AAZ54336	Neisseria meningit
6	86	17.2	1428	21	AAZ54335	Neisseria meningit
7	86	17.2	24158	21	AAAB1532	N. meningitidis po
8	86	17.2	349980	21	AAAF21611	Neisseria meningit
C 9	86	17.2	349980	21	AAAF21612	Neisseria meningit

C 10	86	17.2	837096	21	AA8A1489	N. meningitidis pa
C 11	51.6	10.3	5120	18	AA773218	APx1D gene. Acti
C 12	51.6	10.3	5120	21	AA288585	A. pleuropneumonia
C 13	47.8	9.5	6400	20	AAZ22701	Pseudomonas fluores
C 14	47	9.4	4403765	22	AA199683	Mycobacterium tuberc
C 15	47	9.4	4411529	22	AA199682	Mycobacterium tuberc
C 16	45	9.0	7184	15	AA070050	pxx gene of Pasteur
C 17	43.4	8.7	7183	11	AA006074	Sequence encoding
C 18	43	8.6	24379	18	AA193095	Streptomyces itreno
C 19	43	8.6	24379	19	AAV25925	Streptomyces roseo
C 20	39	7.8	1761	22	AAH66404	C glutamicum codin
C 21	39	7.8	1998	22	AAAF72013	Corynebacterium gl
C 22	39	7.8	349980	22	AAH66358	C. glutamicum codin
C 23	37.6	7.5	1553	23	AA589354	DNA encoding novel
C 24	37.6	7.5	2931	23	AA592554	DNA encoding novel
C 25	37.6	7.5	10732	23	AA510594	Gene encoding a su
C 26	37.6	7.5	4403765	22	AA199683	Mycobacterium tuberc
C 27	37.2	7.4	1194	17	AA585555	Streptomyces prist
C 28	37.2	7.4	4496	17	AA585553	Streptomyces prist
C 29	37.2	7.4	58857	21	AA58471	Nucleotide sequenc
C 30	37	7.4	1951	19	AAV14459	C. acidivorans gam
C 31	37	7.4	1981	21	AAZ50482	Corn sulphate perm
C 32	36.4	7.3	1793	22	AAAD19384	Rat CCAAT/enhancer
C 33	36.2	7.2	16020	21	AAA39283	Streptomyces nogal
C 34	36	7.2	1136	21	AAAC47742	zea mays DNA fragm
C 35	35.8	7.1	836	19	AAAG6157	HSV-2 strain SB5 C
C 36	35.8	7.1	3765	21	AAAC5582	Nucleotide sequenc
C 37	35.8	7.1	57500	21	AAAC55842	Complete nucleotid
C 38	35.8	7.1	117213	19	AAV62176	HSV-2 strain SB5 C
C 39	35.8	7.1	154746	22	AAAD5519	Human herpesvirus
C 40	35.6	7.1	141529	22	AA199682	Mycobacterium tuberc
C 41	35.4	7.1	48300	22	AAAF61281	N. magdaldi bacter
C 42	35	7.0	1119	24	AAAD24963	Human G-protein co
C 43	35	7.0	2025	24	AAAA9036	DNA encoding a hig
C 44	34.8	6.9	1470	21	AAA38391	Pseudomonas sp. WF
C 45	34.8	6.9	1500	22	AAAD19383	Mouse CCAAT/enhanc

ALIGNMENTS	
RESULT	1
AAAT37329/c	AAAT37329 standard; DNA; 28804 BP.
XX	
XX	AAAT37329;
XX	
DT	30-NOV-1996 (first entry)
XX	
DE	Sphingian biosynthetic gene region.
XX	
KW	Sphingian; polysaccharide; spsB gene; glucosyl-IP-transferase
XX	
OS	Sphinomonas strain S88 (ATCC 31554).
XX	
FH	Location/Qualifiers
FT	complement (1942..1944)
FT	/*tag= a
FT	/codon_start=1942..1944
FT	/note="spsG gene putative initiation codon"
FT	complement (3311..3313)
FT	/*tag= b
FT	/codon_start=3311..3313
FT	/note="spsS gene putative initiation codon"
FT	complement (5323..5325)
FT	/*tag= c
FT	/codon_start=5323..5325
FT	/note="spsR gene putative initiation codon"
FT	5526..5528
FT	/*tag= d
FT	/codon_start=5526..5528
FT	/note="spsQ gene putative initiation codon"
FT	complement (7076..7078)
FT	
CDS	

FT	/tag= e
FT	/codon_start= 7076...7078
FT	/note= "spst gene putative initiation codon"
FT	7588...7590
FT	/tag= f
FT	/codon_start= 7588...7590
FT	/note= "spsk gene putative initiation codon"
FT	8643...8645
FT	/tag= g
FT	/codon_start= 8643...8645
FT	/note= "pspl gene putative initiation codon"
FT	complement (10938..10940)
FT	/tag= h
FT	/codon_start= 10938..10940
FT	/note= "pspj gene putative initiation codon"
FT	11569..11571
FT	/tag= i
FT	/codon_start= 11569..11571
FT	/note= "spsf gene putative initiation codon"
FT	12886...12888
FT	/tag= j
FT	/codon_start= 12886..12888
FT	/note= "pspd gene putative initiation codon"
FT	15165...15167
FT	/tag= k
FT	/codon_start= 15165..15167
FT	/note= "spse gene putative initiation codon"
FT	15883..15885
FT	/tag= l
FT	/codon_start= 15883..15885
FT	/note= "urf12 gene putative initiation codon"
FT	16748...16750
FT	/tag= m
FT	/codon_start= 16748..16750
FT	/note= "urf26 gene putative initiation codon"
FT	complement (21082..21084)
FT	/tag= n
FT	/codon_start= 21082..21084
FT	/note= "atrb gene putative initiation codon"
FT	21706...23118
FT	/tag= o
FT	/label= spsb
FT	/product= glucosyl IP-transferase
FT	23238..23240
FT	/tag= p
FT	/codon_start= 23238..23240
FT	/note= "rfsa gene putative initiation codon"
FT	24113...24115
FT	/tag= q
FT	/codon_start= 24113..24115
FT	/note= "rhsc gene putative initiation codon"
FT	24683..24685
FT	/tag= r
FT	/codon_start= 24683..24685
FT	/note= "rhsh gene putative initiation codon"
FT	25744...25746
FT	/tag= s
FT	/codon_start= 25744..25746
FT	/note= "rhsd gene putative initiation codon"
FT	complement (27534..27536)
FT	/tag= t
FT	/codon_start= 27534..27536
FT	/note= "urfi3 gene putative initiation codon"
FT	27747
FT	/tag= u
FT	/codon_start= 27747..27749
FT	/note= "urf34 gene putative initiation codon"
PX	EP728841-A2.
PD	28-AUG-1996.
XX	24-JAN-1996;
XX	96EP-0300467

XX 24-JAN-1995; 95US-0377440.  
PR  
XX  
XX (SHIN-) SHINETSU BIO INC.  
PA (SHE ) SHINETSU CHEM CO LTD.  
XX  
XX Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;  
PI Yamazaki M;  
PI  
XX MPI: 1996-386292/39.  
DR P-PSDB: AAM03997.  
XX  
XX  
XX New isolated DNA from *Sphingomonas* sp. - used for transforming  
PT recipient bacteria to obtain hyper-producers of sphingan  
PT polysaccharide(s).  
XX  
XX  
XX Claim 32; Page 56-70; 105pp; English.  
XX  
XX  
XX A 28.8 kb chromosomal fragment of *Sphingomonas* strain S88 was  
CC isolated on the basis of its ability to restore sphingan  
CC biosynthetic capability to *Sphingomonas* mutant S88m260. It  
CC contains 23-25 genes, including sps genes coding for biosynthesis of  
CC the polysaccharide sphingan, ths genes coding for dTdp-(U)ribose  
CC biosynthesis, atdB genes coding for a transport function and some  
CC unidentified open translation reading frames (unrf). The spsB gene  
CC was identified that is believed to code for glucosyl IP-transferase  
CC (AAM03997), an enzyme catalyzing the first step of assembly of  
CC sphingan carbohydrates. DNA fragments of s88 can be inserted into  
CC a vector in multiple copies and used to produce engineered bacteria  
CC that are hyper-producers of sphingan.  
XX  
XX Sequence 28804 BP; 4974 A; 9806 C; 9228 G; 4796 T; 0 other;  
XX

	Query Match	17.7%;	Score 88.8;	DB 17;	Length 28804;	
	Best Local Similarity	56.5%;	Pred. No. 6,2e-13;			
	Matches 165; Conservative	0;	Mismatches 127;	Indels	0;	Gaps 0.
QY	80	tcttcgcccgcgcgcacctgaattgtcagagacgcccgccatctccaccacggagaactcacg	139			
Dd	18808	TTCCTCCCGGGGCGGTCCAGATCATCAGAACTCCGGTGTCCGCCACCGCGGCCTCACAC	18749			
QY	140	gccgcctctgtctgtccttgttctaactgagccgcttgccgttgccgttgccgttgccgagatc	199			
Dd	18748	GGCGGGTGATGATGTGGCGGGGCTGGCGGATCACACCACCGCTGGCTGGCAATCGGCCGCTG	18689			
QY	200	gacatcgtttgtctcttcacatccagaagaatcgcgcgcggaggaccgctgtaagtgtgtcag	259			
Dd	18688	GAAGTGGTGGCCCCGACGAGGGCCGCAATGCCCCGATGGGGAGACCAAGATCGTGCAG	18629			
QY	260	cgcctcgaagtcgcgcgtgtgtgcgggcacatagtctcgcgatgagccaaacgylcaagyscc	319			
Dd	18628	TCCCCCGAAGAAGCGGTTCTGCTCCGGCCATTCTGTGTGGCGAGGGGCGAGAAGTTCTGAAG	18569			
QY	320	ggcgaattctgatcgaagcttgatccatctgcgcggatgtgtgtgatltygcgc	371			
Dd	18568	GGCAGAGTGTGATCACGCTCGATCCACCATGTCCGACGCGAAGCCC CGC	18517			
<b>RESULT 2</b>						
AAI92474/c						
ID AAI92474 standard; DNA; 28804 BP.						
XX	AAI92474;					
XX	AC					
XX	DT					
XX	04-FEB-1998 (first entry)					
XX	Sphingomonas genus microbe isolated DNA sequence producing sphingan.					
XX	Sphingomonas microbe; sphingan polysaccharide biosynthesis gene;					
KW	sphingan S-88; spsB gene; ss.					
XX						
OS	Sphingomonas sp.					

XX JP09252775-A.  
 XX  
 XX 30-SEP-1997.  
 XX  
 XX 24-JAN-1996; 96JP-0043977.  
 XX  
 XX 16-JAN-1996; 96JP-0004621.  
 XX PR 24-JAN-1995; 95US-0377440.  
 XX  
 XX (SHIN-) SHINETSU BIO INC.  
 XX (SHIE ) SHINETSU CHEM IND CO LTD.  
 XX  
 XX WPI: 1997-530148/49.  
 XX  
 XX Sphingn polysaccharide biosynthesis gene from Sphingomonas microbe  
 PT -can be introduced into Sphingomonas host cells to increase  
 PT sphingn production  
 XX  
 XX Claim 74; Page 34-43; 55pp; Japanese.  
 CC  
 CC The present sequence represents a new DNA sequence which has been  
 CC isolated from a sphingn polysaccharide-producing microbe of the  
 CC genus Sphingomonas. When recombined in a Sphingomonas host, the  
 CC presence of several copies of the sphingn polysaccharide biosynthesis  
 CC gene results in a microbe which expresses high levels of sphingn.  
 CC A method has been developed for producing a Sphingomonas microbe with  
 CC enhanced sphingn polysaccharide expression, by transforming the cell  
 CC with the sphingn polysaccharide biosynthesis gene. The method can  
 CC produce sphingn in large quantities.  
 XX  
 XX Sequence 28804 BP; 4974 A; 9804 C; 9230 G; 4796 T; 0 other;  
 SQ  
 Query Match 17.7%; Score 88.8; DB 18; Length 28804;  
 Best Local Similarity 56.5%; Pred. No. 6.2e-13;  
 Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;  
 QY 80 ttcttcgagcgccgcctcgaattgtcgagacgcgcacatctccacccgagactcacg 139  
 Db 18808 TTCTGTCGCGCGCGCGCTCGAGATCATCGAAGCTCGGTTGTCCGCCCGCGCCCTCCACC 18749  
 QY 140 gccgcctgtgtgtgccttgttctactcgcgcgtgagcgtggcggtgtcgagcagatc 199  
 Db 18748 GCCCGGGTGATGTGTGGCGGCGTGGCGATCACACCGCTGCGTGGCAATCGGCGCGTGG 18689  
 QY 200 gacatcgctgtcttcgcataccagaagaatcgctgcgcggcgagacgcgtgtaagctgttcag 259  
 Db 18688 GAAGTGTGTGGCGCGCCACGACGAGGCGCGCATCGCCCGATGCGCGAGACCAAGATCTGTGACG 18629  
 QY 260 ccgctcgaggttcgagcgtgtgtgtcgggccacatcgttcgcgcgcacaaacgctcagagcc 319  
 Db 18628 TCCCGCCGAAGGCGGTATCGTCCGGCGCATTCGTGTGGGGGAGGAGGCGACGAAGGTTCAGAG 18569  
 QY 320 ggcgcgatttcgatcgagctgatcatcattcgcgggtgtgtgtgaagtgtgcgc 371  
 Db 18568 GCGCGAGTGTGTATCATCGCTCGATCCACCATGTGTGGCAGCCGAAGCGCGCG 18517  
 RESULT 3  
 AA99812/c  
 ID AA99812 standard; DNA; 28804 BP.  
 XX  
 XX AA99812;  
 XX  
 XX 14-JUN-1999 (first entry)  
 XX  
 XX Sphingomonas S88 sps gene cluster.  
 DE  
 DE Xanthan gum; gum; exopolysaccharide; gum gene cluster;  
 KW recombination; Xanthomonas; Sphingomonas; sphingn S88; sugar;  
 XX substrate; lactose; sucrose; starch; ss.  
 XX

OS	Sphingomonas sp. S88.
XX	
PN	WO9869642-A1.
XX	
PD	17-DEC-1998.
XX	
PF	12-JUN-1998; 98WO-US12322.
XX	
PR	12-JUN-1998; 98US-0096942.
PR	12-JUN-1997; 97US-0049428.
PR	11-JUN-1998; 98US-0096867.
PA	(SHIN-) SHINETSU BIO INC.
PA	(SHITE) SHINETSU CHEM CO LTD.
XX	
PI	Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;
PI	Tamazaki M;
DR	WPI; 1999-080915/07.
XX	
PT	Production of exopolysaccharide, e.g. xanthan gum by a Sphingomonas
PT	species bacterium - containing the gumb-X genes of X. campestris,
PT	and using lactose as a C source
XX	
XX	Example 1; Page 36-46; 50pp; English.
XX	
CC	Xanthan gum is an exopolysaccharide produced by Xanthomonas species
CC	and encoded by the gum gene cluster. The gum gene cluster of
CC	Xanthomonas campestris can be inserted into a Sphingomonas species
CC	of bacterium. The gum gene cluster is inserted into the Sphingomonas
CC	S88 sps gene cluster which itself expresses an exopolysaccharide,
CC	sphingane S88. The new novel recombinant Sphingomonas strain can then
CC	produce xanthan gum from sugar substrates i.e. the cheese-making
CC	by-product whey lactose, sucrose or starch. These are substrates
CC	which the Xanthomonas campestris bacterium cannot utilise in its
CC	production of Xanthan gum. Use of the Sphingomonas species also
CC	minimises the presence of contaminating cellulases in the growth
CC	medium into which the xanthan gum is secreted.
XX	
SQ	Sequence 28804 BP; 4974 A; 9804 C; 9230 G; 4796 T; 0 other;
<hr/>	
Query Match	17.7%; Score 88.8; DB 20; Length 28804;
Best Local Similarity	56.5%; Pred. No. 6,2e-13;
Matches 165; Conservative	0; Mismatches 127; Indels 0; Gaps 0;
QY	80 ttctgcgcgccgccacctgaattgtcggagacgcgcacatcccaaccgcgagactcacg 139
DB	18808 TTCCTGCCGCGCGGTCTGCAGATCATTCAGACACTCCCGTGTCCGCCACCGCGCCTCAC 18749
QY	140 gcgcgcttgcttgcgttccttgtlctaalcigcgcgttgcgttggcggtctcgcgcagatc 199
DB	18748 GGCCGGGTTATGTTGGCGGGGCTGGCGATCACCGCGCTGTGGCTGGCAATCGCGCGCGTG 18689
QY	200 gcatcgtttgcttcgcgatccagaagaactcgttcgcgcggaggacgcgtaaacgttggtcaa 259
DB	18688 GAATGTGTGTGCCCCCAGCAGAGGGCGCGCATCGCCCAGATGCGGAGACCAAGAATCGTGCAG 18629
QY	260 ccgctcgaagtctgcgcgttgttgcgcacaatactgtccgcatgcatgcatgcacaacgctcaagcc 319
DB	18628 TCCCCCGAAGAGGGTATCGTCCGGCGCATTTGHTGGGCGGAGGGGCAAGAAGTTTCAGAAAG 18569
QY	320 gccgcgattcttatcgagcttgatccatctcgcggagtgtgtgtgatgttgcgc 371
DB	18568 GCCAGAGTGTGATCACGCTCGATCCACCACCATGTGGGAGCCGAAGCCGCCG 18517
<hr/>	
RESULT 4	
AAVB1474/C	
ID AAVB1474 standard; DNA; 28804 BP.	
XX	
AC AAVB1474;	
XX	





CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent

PI Frazier CM, Hickey E, Peterson J, Tettelin H, Venter JC; Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato

PI Rappuoli R, Pizza M;



```

Db 323706 TCCTGTTGCGCAAAATGATATTGTGGCGGCTTCGGGCAAAACGGTGTCTCGGCGG 323647
QY 242 cgtgaagctgcttcagcgcctcgagctggtgctcgagccactatgtccgcgat 301
Db 323646 CGCAGCAAAACCATCATCCAGCCGCTGGAAACGGCGGTGTTAAAGCGGTACATGTGCCGCGAC 323587
QY 302 ggcacaaccgtcaagccgagatctctgacagctgcatccatctcgagtggtg 361
Db 323586 GGGCAGCATGTGAAACAGGAGAAACGCTGGCGGAACCTGGAGCTGTGGGAACAGACAGC 323527
QY 362 gatgttgccg 371
Db 323526 GATGTGTGC 323517

RESULT 9
AAF21612/c
ID AAF21612 standard; DNA; 349980 BP.
AAF21612;
AA 13-MAR-2001 (first entry)
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX ds.
XX Neisseria meningitidis.
XX MO20006791-A1.
XX 09-NOV-2000.
XX 08-MAR-2000; 2000WO-US05928.
XX 30-APR-1999; 99US-0132068.
XX 08-OCT-1999; 99WO-US23573.
XX 28-FEB-2000; 2000GB-0004695.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Pizze M, Hickey E, Peterson J, Tettelin H, Venter JC, Masiagnani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX WPI: 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
XX frames are used to detect, treat and prevent Neisserial infections -
XX
XX Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of
XX Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
XX to AAF21613 represent fragments of the NMB genomic sequence, as the
XX sequence was too long to go in a record on its own it was split into 8
XX sequences which overlap each other at the beginning and end of each
XX sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
XX the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
XX the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
XX Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
XX AAF21606 represent PCR primers which are used in the exemplification of
XX the present invention. The NMB genome and fragments from it have
XX antibacterial activity, and can be used in vaccines and gene therapy.
XX Neisseria nucleic acids, proteins and/or antibodies which binds to the
XX proteins can be used in compositions for treating or preventing infection
XX due to Neisserial bacteria or as a diagnostic reagent for detecting the
XX presence of Neisserial bacteria or of antibodies raised to Neisserial
XX bacteria. Computers, computer memory, computer storage medium or computer

```

```

CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
XX SQ Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;
XX
XX Query Match 17.2%; Score 86; DB 21; Length 349980;
XX Best Local Similarity 54.8%; Pred. No. 4.8e-12;
XX Matches 170; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
XX
QY 62 acccggaacgagctggcctctccgagcgccgcctgcaaatgtcgagagcgcccatct 121
Db 23826 ACGGCGGAAGAAACAGGCGCTTTTGGCCGCAATTGGAAGTGAACGCGGCTCT 23767
QY 122 ccacacgcgagatcacgagccgctgtctgtctgtcttactacgcgcgtg 181
Db 23766 GCCGCTCCGAAATGGCGCGCGGCTTTATTATGGCGCTTTGGCTTTGTGTGG 23707
QY 182 gcgagctcgcgagagatgcacatcgttctgtcatccagaagatcgccggcgac 241
Db 23705 TCCTGTTGCGCAAAATGATATTGTGGCGGCTTCGGCAAAACGGTGTCTCGGCGG 23647
QY 242 cgtgaagctgcttcagcgcctcgagctggtgctcgagccactatgtccgcgat 301
Db 23646 CGCAGCAAAACCATCATCCAGCCGCTGGAAACGGCGGTGTTAAAGCGGTACATGTGCCGCGAC 23587
QY 302 ggcacaaccgtcaagccgagatctctgacagctgcatccatctcgagtggtg 361
Db 23586 GGGCAGCATGTGAAACAGGAGAAACGCTGGCGGAACCTGGAGCTGTGGGAACAGACAGC 23527
QY 362 gatgttgccg 371
Db 23526 GATGTGTGC 23517

RESULT 10
AAA81489/c
ID AAA81489 standard; DNA; 837096 BP.
AAA81489;
AA 04-DEC-2000 (first entry)
DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; Memb; ds.
XX
XX Neisseria meningitidis.
XX MO200022430-A2.
XX
XX 20-APR-2000.
XX 08-OCT-1999; 99WO-US23573.
XX 09-OCT-1998; 98US-0103794.
XX 30-APR-1999; 99US-0132068.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masiagnani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizze M;
XX WPI: 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and

```



DB 2831 ggaatgaatgaatgttaccgacacactagagcgtgatgaaacccggtgtcaaaaaa 2890  
 QY 127 cggagagactcaagcgcgccttgctgtcgtctgtctactgcccgtgcggtggggg 186  
 DB 2831 accgagattgacgcgtattatgattgctgtccatttttgccattagttat 2950  
 QY 187 tctcggcaagatcagatctgtctctcgtcatccaaagatcgtgcgcgcacgtgt 246  
 DB 2951 tgcagtcacgtagaatgtgtgacccgaaggttaattagcgttttgacgcgtag 3010  
 QY 247 aaagctgttcacgcgcgtcgaaggtgcggtgtgcggccacatcgtccgcgtagcca 306  
 DB 3011 caaagaatlaagccgattgaaacgccttggttaagaatcttggtgcaagacgaca 3070  
 QY 307 aacgtcaagcgcgcgcgtatctcgtacgcgtgacatccattgcgggtgtgtgcat 364  
 DB 3071 attgttgagaagaatcagttgctgttacaactgacccgattggagcgcgtgcgat 3128

# BLAST 12

AA288585  
 ID AA288585 standard; DNA; 5120 BP.

AA288585;

05-MAY-2000 (first entry)

A. pleuropneumoniae DNA encoding clyIB and clyID.

RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;  
 antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;  
 pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;  
 shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;  
 urinary infection; peritonitis; meningitis; gastroenteritis;  
 passive immunization; clyIB; clyID; ds.

Actinobacillus pleuropneumoniae.

Key Location/Qualifiers

CDS 592..2715

/\*tag= a

/product= "clyIB"

2724..4160

/\*tag= b

/product= "clyID"

US6019984-A.

01-FEB-2000.

23-DEC-1996; 96US-0772270.

01-MAR-1995; 95US-0396244.

(UYGU-) UNIV GUELPH.

Mallard B, Rosendal S, MacInnes J, Ricciatti P;

WP1: 2000-146864/13.

P-PSDB: AAY51407, AAY51408.

Bacterial preparation comprising microorganisms which produce a member  
 of the Repeats in Toxins (RTX) family, useful for treating swine  
 pleuropneumonia, arthritis in swine, shipping fever and abortion in  
 cattle, and sleepy foal disease -

Disclosure: Column 61-66; 96pp; English.

This invention describes a novel bacterial preparation (I) which  
 comprises one or more isolated and purified strain(s) of a microorganism,  
 cultured in tryptone yeast extract (TYE) broth, which produces one or  
 more RTX toxins (belonging to the family of toxins referred to as Repeats

in Toxins), where the strain(s) have at least one RTX toxin which is  
 cell-associated. The products of the invention have immunostimulatory,  
 antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.  
 The bacterial preparation may be used as vaccines for the prophylaxis and  
 treatment of infectious diseases caused by strains of microorganisms  
 which produce one or more RTX toxins. The infectious diseases are swine  
 pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;  
 septicemia, nephritis, endocarditis and arthritis in piglets; shipping  
 fever and abortion in cattle; whooping cough, sleepy foal disease or  
 joint ill (purulent nephritis, arthritis) in foals; septicemia,  
 polyarthritis and abortion in horses; and urinary infections,  
 peritonitis, meningitis, and gastroenteritis. The bacterial preparations  
 may also be used to prepare antibodies which may be used as a means of  
 passive immunization. This sequence encodes the Actinobacillus  
 pleuropneumoniae clyIB and clyID proteins described in the method of  
 the invention.

Sequence 5120 BP; 1580 A; 880 C; 1139 G; 1521 T; 0 other;

Query Match 10.3%; Score 51.6; DB 21; Length 5120;  
 Best Local Similarity 48.3%; Pred. No. 0.0012;  
 Matches 144; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 67 ggaagctgctgcctctcgcgcgcgcgcgcgaatgtcgaagcgcgcacatctccac 126  
 DB 2831 ggaatgaatgaatgttaccgacacactagagcgtgatgaaacccggtgtcaaaaaa 2890  
 QY 127 cggagagactcaagcgcgccttgctgtcgtctgtctactgcccgtgcggtggggg 186  
 DB 2831 accgagattgacgcgtattatgattgctgtccatttttgccattagttat 2950  
 QY 187 tctcggcaagatcagatctgtctctcgtcatccaaagatcgtgcgcgcacgtgt 246  
 DB 2951 tgcagtcacgtagaatgtgtgacccgaaggttaattagcgttttgacgcgtag 3010  
 QY 247 aaagctgttcacgcgcgtcgaaggtgcggtgtgcggccacatcgtccgcgtagcca 306  
 DB 3011 caaagaatlaagccgattgaaacgccttggttaagaatcttggtgcaagacgaca 3070  
 QY 307 aacgtcaagcgcgcgcgtatctcgtacgcgtgacatccattgcgggtgtgtgcat 364  
 DB 3071 attgttgagaagaatcagttgctgttacaactgacccgattggagcgcgtgcgat 3128

## RESULT 13

AA222701  
 ID AA222701 standard; DNA; 6400 BP.

AA222701;

15-FEB-2000 (first entry)

Pseudomonas fluorescens ABC transporter cassette operon II.

ATP-binding cassette; transporter; operon; LipCD; Serratia marcescens;

microbe; protein secretion; ds.

Pseudomonas fluorescens.

Key Location/Qualifiers

CDS 278..2020

/\*tag= a

/product= "component A"

2017..3348

/\*tag= b

/product= "component B"

3354..4670

/\*tag= c

/product= "component C"

JP11276172-A.









...

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2002, 22:46:27 ; Search time 43.46 Seconds  
(without alignments)  
2831.624 Million cell updates/sec

Title: US-09-484-577A-3  
Perfect score: 501  
Sequence: 1 actctccagctcaccga.....cagaatgcygcatgatcat 501

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

al number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/lna/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/lna/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/lna/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/lna/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/lna/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/lna/Dackfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	88.8	17.7	28804	2	US-08-592-874-1
C 2	88.8	17.7	28804	3	US-09-096-942-2
C 3	88.8	17.7	28804	3	US-09-096-867-2
C 4	51.6	10.3	5120	3	US-08-772-270A-6
C 5	51.6	10.3	8370	2	US-08-488-706-1
C 6	47	9.4	4403765	4	US-09-103-840A-2
C 7	47	9.4	4411529	4	US-09-103-840A-1
C 8	37.6	7.5	4403765	4	US-09-103-840A-2
C 9	37.2	7.4	1194	4	US-08-765-907A-9
C 10	37.2	7.4	4496	4	US-08-765-907A-6
C 11	37	7.4	1951	3	US-08-922-865-1
C 12	35.6	7.1	4411529	4	US-09-103-840A-1
C 13	34.8	6.9	1500	4	US-09-593-711A-10
C 14	34.6	6.9	8438	1	US-07-945-283-1
C 15	34.4	6.9	1879	3	US-08-403-852D-5
C 16	34.4	6.9	1879	3	US-08-510-646B-5
C 17	34.4	6.9	1879	3	US-09-231-818-5
C 18	34.2	6.8	10095	3	US-08-822-586-45
C 19	34	6.8	1758	1	US-08-078-222B-1
C 20	34	6.8	1758	2	US-08-661-330A-1
C 21	34	6.8	1758	3	US-09-038-217A-1
C 22	34	6.8	49272	1	US-08-614-770A-1
C 23	33.8	6.7	933	4	US-09-105-390-43
C 24	33.8	6.7	1008	4	US-09-105-390-59
C 25	33.8	6.7	2810	4	US-09-105-390-6
C 26	33.2	6.6	2293	4	US-08-604-913B-12
C 27	33.2	6.6	2787	4	US-09-105-537-40

C 28	33.2	6.6	3004	1	US-08-276-213-6	Sequence 6, Appl
C 29	33.2	6.6	5970	3	US-09-320-878-21	Sequence 21, Appl
C 30	33	6.6	789	4	US-09-164-193-6	Sequence 6, Appl
C 31	33	6.6	1016	4	US-09-164-193-4	Sequence 4, Appl
C 32	32.8	6.5	5247	1	US-08-920-812-15	Sequence 15, Appl
C 33	32.8	6.5	5247	1	US-08-920-827-15	Sequence 15, Appl
C 34	32.8	6.5	5247	1	US-08-921-177-15	Sequence 15, Appl
C 35	32.8	6.5	5247	1	US-08-362-577C-15	Sequence 15, Appl
C 36	32.8	6.5	5247	2	US-08-920-828-15	Sequence 15, Appl
C 37	32.6	6.5	3576	4	US-08-976-259-79	Sequence 79, Appl
C 38	32.6	6.5	23673	4	US-09-773-816-1	Sequence 129, App
C 39	32.4	6.5	3065	4	US-08-599-895-1	Sequence 1, Appl
C 40	32.4	6.5	3877	2	US-08-599-895-1	Sequence 1, Appl
C 41	32.4	6.5	3877	3	US-09-211-290-1	Sequence 1, Appl
C 42	32.4	6.5	3877	3	US-09-322-676-1	Sequence 1, Appl
C 43	32.4	6.5	3877	4	US-09-466-036A-1	Sequence 1, Appl
C 44	32.2	6.4	295	4	US-09-082-092-2	Sequence 2, Appl
C 45	32.2	6.4	720	4	US-08-998-416-662	Sequence 662, App

## ALIGNMENTS

RESULT 1  
US-08-592-874-1/C  
Sequence 1, Application US/08592874  
Patent No. 5854034  
GENERAL INFORMATION:  
APPLICANT: POLLOCK, THOMAS J.  
APPLICANT: YAMAZAKI, MOTOHIDE  
APPLICANT: THORNE, LINDA  
APPLICANT: MIKOLAJCZAK, MARCIA  
APPLICANT: ARMENTROUT, RICHARD W.  
TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JULES E. GOLDBERG  
STREET: 261 MADISON AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,874  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/377,440  
FILING DATE: 24-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: GOLDBERG, JULES E.  
REGISTRATION NUMBER: 24,408  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-986-4090  
TELEFAX: 212-818-9479  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28804 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FRAGMENT TYPE: N-terminal  
US-08-592-874-1  
Query Match 17.7%; Score 88.8; DB 2; Length 28804;



MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/772,270A  
 FILING DATE: December 23, 1996  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gravelle, Michelle  
 REGISTRATION NUMBER: 40,261  
 REFERENCE/DOCKET NUMBER: 6580-81  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5120 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: Single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGIN: SOURCE:  
 ORGANISM: *Actinobacillus pleuropneumoniae*

Query Match	10.3%	DB 3:	Length	5120;
Best Local Similarity	48.3%	Pred. No.	7.3e-05;	
Matches 144;	Conservative	0;	Mismatches	154;
			Indels	0;
			Gaps	0;

[illegible]

```

1 RESULT 5
2 US-08-488-706-1
3 Sequence 1, Application US/08488706
4 Patent No. 5994525
5
6 GENERAL INFORMATION:
7
8 APPLICANT: Kamp, Elbarte M.
9 APPLICANT: Smits, Marinus A.
10 TITLE OF INVENTION: Recombinant Vaccine For Prevention and/or
11 TITLE OF INVENTION: Treatment of Pleuropneumonia Infections
12 NUMBER OF SEQUENCES: 3
13
14 CORRESPONDENCE ADDRESSES:
15 ADDRESSEE: Adelman, Frayne & Schwab
16 STREET: 708 Third Avenue
17 CITY: New York
18 STATE: New York
19 COUNTRY: USA
20 ZIP: 10017
21
22 COMPUTER READABLE FORM:
23 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
24 COMPUTER: IBM PC or compatible
25 OPERATING SYSTEM: DOS 3.31
26
27
28
29
30

```

```

: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/488,706
: FILING DATE: 09-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/722,971
: FILING DATE: 28-JUNE-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Cinnamon, Jay S.
: REGISTRATION NUMBER: 24,156
: REFERENCE/DOCKET NUMBER: 201,875
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-949-9022
: TELEFAX: 212-949-9190
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 8370
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
:
: US-08-488-706-1

```

Query Match	10.3%;	Score 51.6;	DB 2;	Length 8370;
Best Local Similarity	48.3%;	Pred. No. 8.3e-05;		
Matches 144;	Conservative	0;	Mismatches 154;	Indels 0;
				Gaps 0;

QY	67	gaacagagctgagccctccctgcgcgcgcgcctccgaattgtcgaagacgcgcgcattccac	126
DB	6081	GGATTGAAATGAAATTTTACCCTGCACACTTAGAGCTATTGAAACACCGGTGTA	6140
QY	127	cgcagacacacagcgcgcctctgctctgctctgcttctactgcgcgcgtgcgttgcgcgcgc	186
DB	6141	ACCGAATTGATGCGCTATTATTAAATTATGCTGTTCCATTATTGTCATTAGTATTTCAT	6200
QY	187	tctcgcacagatctgacatcgtttgcttcttcgtatccagaagaatcgtgcgcgcgcgcctgt	246
DB	6201	TGTCACGCACGTAAGAAATTTGTGGCAGCCGACACGGTAATTAACGTTTACGCGACCGTAG	6260
QY	247	aaagctcgtctcagcgcgcctcgaagtcgcgcgttgctgcgcgcgcacatctcgcgcattgcga	306
DB	6261	CANAAGAAATTAAAGCCGATTGAAACGCGCTTTGTTAAAGAAATCTTTGTCGAAGCGGACA	6320
QY	307	aaccgctcaagc	364
DB	6321	ATTGTTGGAAGAAAGATCAGTTGCTGTTACACTTAACCGCATTTGGAGCCGATTCGCGAT	6378

```

RESULT 6
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ. ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g

```



[illegible]

RESULT 10  
 US-08-765-907A-6  
 Sequence 6, Application US/08765907A  
 Patent No. 6352839  
 GENERAL INFORMATION:  
 APPLICANT: BLANC, Veronique  
 APPLICANT: THIBAUT, Denis  
 APPLICANT: BAMA-S-JACQUES, Nathalie  
 APPLICANT: BLANCHE, Francis  
 APPLICANT: COUZET, Joel  
 APPLICANT: BARRIERE, Jean-Claude  
 APPLICANT: DEBUSCHE, Laurent  
 APPLICANT: FAMECHON, Alain  
 APPLICANT: PARIS, Jean-Marc  
 APPLICANT: DUTRUC-ROSETT, Gilles  
 TITLE OF INVENTION: Streptogramins And Method For Preparing Same By  
 TITLE OF INVENTION: Mutasynthesis  
 FILE REFERENCE: Streptogramin genes  
 CURRENT APPLICATION NUMBER: US/08/765, 907A  
 CURRENT FILING DATE: 1997-03-20  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 6  
 LENGTH: 4496  
 TYPE: DNA  
 ORGANISM: Streptomyces pristinaespiralis  
 US-08-765-907A-6

	Query Match	7.4%	Score 37.2:	DB 4:	Length 4496;
	Best Local Similarity	44.7%:	Pred. No. 0.47:		
	Matches 144:	Conservative	0:	Mismatches 178:	Indels 0:
	Gaps				
QY	68	gacgaagctgcgtccttcctccggcgccgcccctcgaaattgtcggagaagccgccatctccacc	127		
Db	2286	gccccgccgctcgagacatgactgcggagacagacgcgcagagactaccacgcttcgtc	2345		
QY	128	gcgagaccacacgcgcgcgttcttgtagcttgtttactgcgcgctggcggttggcgsgyt	187		
Db	2346	aagaagaacctatgcgcctatgcaagcctcatctccaagacaagctcacgcagcgcgcgc	2405		
QY	188	ctcgacagatcgacacatcgtttcttcgtatccagaagaacgcgtgcggcgagacgtgta	247		
Db	2406	cgcgccgcgaaagacactgatctccgcactcgtcgcgcgcgctggaggcggaagacac	2465		
QY	248	aagctggttcacgcgctcgaggttcggcggtggcgcgccactatgtctcgcgatgtgcaa	307		
Db	2466	accgaagcagacatctgcctaatctcggcgctctctcgtatggccgcgcacgttccacc	2525		

[illegible]

```

1      RESULT 11
2      US-08-922-865-1/c
3      : Sequence 1, Application US/08922865
4      : Patent No. 6090616
5      : GENERAL INFORMATION:
6      : APPLICANT:
7      : TITLE OF INVENTION: MICROORGANISM, LACTAMASE ENZYME OBTAINED
8      : TITLE OF INVENTION: THEREFROM, AND THEIR USE
9      : NUMBER OF SEQUENCES: 2
10     : COMPUTER READABLE FORM:
11     : MEDIUM TYPE: Floppy disk
12     : COMPUTER: IBM PC compatible
13     : OPERATING SYSTEM: PC-DOS/MS-DOS
14     : SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
15     : CURRENT APPLICATION NUMBER: US/08-922, 865
16     : APPLICATION NUMBER: US/08-922, 865
17     : FILING DATE:
18     : INFORMATION FOR SEQ ID NO: 1:
19     : SEQUENCE CHARACTERISTICS:
20     : LENGTH: 1951 base pairs
21     : TYPE: nucleic acid
22     : STRANDEDNESS: single
23     : TOPOLOGY: linear
24     : MOLECULE TYPE: DNA (genomic)
25     : HYPOTHETICAL: NO
26     : ANTI-SENSE: NO
27     : ORIGINAL SOURCE:
28     : ORGANISM: Comamonas acidovorans
29     : FEATURE:
30     : NAME/KEY: CDS
31     : LOCATION: 49..1773
32     : US-08-922-865-1

```

	Query Match	7.48:	Score 37:	DB 3:	Length 1951,
	Best Local Similarity	53.9%:	Pred. No. 0.43:		
	Matches	76:	Conservative	0:	Mismatches 65; Indels 0; Gaps 0.
QY	228	tctgtccgagcgacacctgttaaaagtcygtltcaagcgcctcgagtgctggcgtgtgtcgaggcca	287		
Db	434	TCTGTCCGCGGCGGCATGTCACAGCAAGTCCACCACAGCAGGTCTGGCGGGCCCC	375		
QY	288	cctatgttcgcgatgtgcccacaacgcgtcaaagccgcgagattcttgaatcgagtcgtgacat	347		
Db	374	TTTCAGGCCACAGGGGCGCGACAGAGAAGTGAGAGCGTGGACATCATCACGTCGGCGCAGTCT	315		
QY	348	tccgcgagtgtgtgtgtagttgtg	368		
Db	314	TCCGCGCTGTCTGCTCTTG	294		

```

RESULT 12
US-09-103-840A-1/C
: Sequence 1, Application US/09103840A
: Patent No. 6294328
:
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
:
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
:
: TITLE OF INVENTION: TUBERCULOSIS
:
: FILE REFERENCE: 24366-20007.00

```









OM of: US-09-484-577a-4 to: GenBank: \* out\_format: pfs  
 Date: Jul 16, 2002 2:10 AM

About: Results were produced by the Gencore software, version 4.5,  
 Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=firmer+p2n.model -DEV=xlh  
 -O=cpn2.1/USPTO.spool/US09484577/runat.15072002.153752.8137/app-query.fasta\_1.184  
 -DB=GenBank -GPM=fastap -SUFFIX=rg -GAPOP=12.000 -GAPEXT=4.000  
 -MIMMARCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500  
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
 -DELEXT=7.000 -START=1 -MATRIX=bloms62 -TRANS=human40.cdi  
 -LIST=45 -DOCALL=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0  
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500  
 -MINLEN=0 -MAXLEN=200000000 -USER=US09484577@CGN1.1\_3836  
 -NCPU=6 -ICPU=3 -LONGLOG -DEV=TIMEOUT=120 -WARN\_TIMEOUT=30  
 -NO\_XLPHY\_WAIT -THREADS=1

## Search information block:

Query: US-09-484-577a-4  
 Query length: 124  
 Database: GenBank: \*  
 Database sequences: 1797656  
 Database length: 187333701  
 Search time (sec): 1788.330000

## score\_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_Ba:AF414151	+	243.00	370.16	3.0e-12	334050	Yersinia pestis str
gb_Ba:AE004049	+	241.00	394.27	1.3e-13	15552	AE004049 Yersinia pestis str
gb_Ba:NME391263	+	238.00	410.48	1.7e-14	1428	AE004049 Yersinia pestis str
gb_Ba:AF121772	+	238.00	400.63	5.9e-13	1428	AE004049 Yersinia pestis str
gb_Ba:AE002524	-	238.00	391.76	1.9e-13	11381	AE002524 Neisseria meningitidis
gb_Pat:AX044033	-	238.00	360.85	9.8e-12	349980	AX044033 Sequence 112 from B
gb_Pat:AX044034	-	238.00	360.85	9.8e-12	349980	AX044034 Sequence 113 from B
gb_Ba:NME391260	+	235.00	389.80	2.4e-13	7824	AE004049 Yersinia pestis str
gb_Ba:NME391260	+	235.00	356.15	1.8e-11	326301	AE004049 Yersinia pestis str
gb_Ba:AL646077	-	218.00	330.84	4.6e-10	189050	AL646077 Ralstonia solanaceae
gb_Ba:SSU51197	-	215.00	342.47	1.0e-10	28804	U51197 Sphingomonas 588 sph
gb_Pat:AR068625	-	205.00	343.44	8.9e-11	3526	AR068625 Sequence 1 from Pat
gb_Ba:AE273216	+	182.00	296.61	3.7e-08	6952	AE273216 Rhizobium leguminosar
gb_Ba:AF314506	+	182.00	296.61	3.7e-08	6952	AF314506 Mannheimia hemolytic
gb_Ba:AF314508	+	182.00	296.61	3.7e-08	6952	AF314508 Mannheimia hemolytic
gb_Ba:AF314510	+	182.00	296.61	3.7e-08	6952	AF314510 Mannheimia hemolytic
gb_Ba:AF314512	+	182.00	296.61	3.7e-08	6952	AF314512 Mannheimia hemolytic
gb_Ba:AF314514	+	182.00	296.61	3.7e-08	6952	AF314514 Mannheimia hemolytic
gb_Ba:AF314515	+	182.00	296.61	3.7e-08	6952	AF314515 Mannheimia hemolytic
gb_Ba:AF314517	+	182.00	296.61	3.7e-08	6952	AF314517 Mannheimia hemolytic
gb_Ba:AF314519	+	182.00	296.61	3.7e-08	6952	AF314519 Mannheimia hemolytic
gb_Ba:AF314521	+	182.00	296.61	3.7e-08	6952	AF314521 Mannheimia hemolytic
gb_Ba:AF314523	+	182.00	296.61	3.7e-08	6952	AF314523 Mannheimia hemolytic
gb_Ba:AF314524	+	182.00	296.61	3.7e-08	6952	AF314524 Mannheimia hemolytic
gb_Ba:AF314526	+	182.00	296.61	3.7e-08	6952	AF314526 Mannheimia hemolytic
gb_Ba:PA314527	+	181.00	295.48	4.3e-08	7742	AF314527 Mannheimia hemolytic
gb_Ba:AF314529	+	181.00	294.83	4.6e-08	6952	AF314529 Mannheimia hemolytic
gb_Ba:AF314531	+	181.00	294.83	4.6e-08	6952	AF314531 Mannheimia hemolytic
gb_Ba:AF314533	+	181.00	294.83	4.6e-08	6952	AF314533 Mannheimia hemolytic
gb_Ba:AF314535	+	181.00	294.83	4.6e-08	6952	AF314535 Mannheimia hemolytic
gb_Ba:AF314537	+	181.00	294.83	4.6e-08	6952	AF314537 Mannheimia hemolytic
gb_Ba:AF314539	+	181.00	294.83	4.6e-08	6952	AF314539 Mannheimia hemolytic
gb_Ba:AF314541	+	181.00	294.83	4.6e-08	6952	AF314541 Mannheimia hemolytic
gb_Ba:AF314543	+	181.00	294.83	4.6e-08	6952	AF314543 Mannheimia hemolytic
gb_Ba:AF314545	+	181.00	294.83	4.6e-08	6952	AF314545 Mannheimia hemolytic
gb_Ba:AF314547	+	181.00	294.83	4.6e-08	6952	AF314547 Mannheimia hemolytic
gb_Ba:AF314549	+	181.00	294.83	4.6e-08	6952	AF314549 Mannheimia hemolytic
gb_Ba:AF314551	+	181.00	294.83	4.6e-08	6952	AF314551 Mannheimia hemolytic
gb_Ba:AF314553	+	181.00	294.83	4.6e-08	6952	AF314553 Mannheimia hemolytic
gb_Ba:AF314555	+	181.00	294.83	4.6e-08	6952	AF314555 Mannheimia hemolytic
gb_Ba:AF314557	+	181.00	294.83	4.6e-08	6952	AF314557 Mannheimia hemolytic
gb_Ba:AF314559	+	181.00	294.83	4.6e-08	6952	AF314559 Mannheimia hemolytic
gb_Ba:AF314561	+	181.00	294.83	4.6e-08	6952	AF314561 Mannheimia hemolytic
gb_Ba:AF314563	+	181.00	294.83	4.6e-08	6952	AF314563 Mannheimia hemolytic
gb_Ba:AF314565	+	181.00	294.83	4.6e-08	6952	AF314565 Mannheimia hemolytic
gb_Ba:AF314567	+	181.00	294.83	4.6e-08	6952	AF314567 Mannheimia hemolytic
gb_Ba:AF314569	+	181.00	294.83	4.6e-08	6952	AF314569 Mannheimia hemolytic
gb_Ba:AF314571	+	181.00	294.83	4.6e-08	6952	AF314571 Mannheimia hemolytic
gb_Ba:AF314573	+	181.00	294.83	4.6e-08	6952	AF314573 Mannheimia hemolytic
gb_Ba:AF314575	+	181.00	294.83	4.6e-08	6952	AF314575 Mannheimia hemolytic
gb_Ba:AF314577	+	181.00	294.83	4.6e-08	6952	AF314577 Mannheimia hemolytic
gb_Ba:AF314579	+	181.00	294.83	4.6e-08	6952	AF314579 Mannheimia hemolytic
gb_Ba:AF314581	+	181.00	294.83	4.6e-08	6952	AF314581 Mannheimia hemolytic
gb_Ba:AF314583	+	181.00	294.83	4.6e-08	6952	AF314583 Mannheimia hemolytic
gb_Ba:AF314585	+	181.00	294.83	4.6e-08	6952	AF314585 Mannheimia hemolytic
gb_Ba:AF314587	+	181.00	294.83	4.6e-08	6952	AF314587 Mannheimia hemolytic
gb_Ba:AF314589	+	181.00	294.83	4.6e-08	6952	AF314589 Mannheimia hemolytic
gb_Ba:AF314591	+	181.00	294.83	4.6e-08	6952	AF314591 Mannheimia hemolytic
gb_Ba:AF314593	+	181.00	294.83	4.6e-08	6952	AF314593 Mannheimia hemolytic
gb_Ba:AF314595	+	181.00	294.83	4.6e-08	6952	AF314595 Mannheimia hemolytic
gb_Ba:AF314597	+	181.00	294.83	4.6e-08	6952	AF314597 Mannheimia hemolytic
gb_Ba:AF314599	+	181.00	294.83	4.6e-08	6952	AF314599 Mannheimia hemolytic
gb_Ba:AF314601	+	181.00	294.83	4.6e-08	6952	AF314601 Mannheimia hemolytic
gb_Ba:AF314603	+	181.00	294.83	4.6e-08	6952	AF314603 Mannheimia hemolytic
gb_Ba:AF314605	+	181.00	294.83	4.6e-08	6952	AF314605 Mannheimia hemolytic
gb_Ba:AF314607	+	181.00	294.83	4.6e-08	6952	AF314607 Mannheimia hemolytic
gb_Ba:AF314609	+	181.00	294.83	4.6e-08	6952	AF314609 Mannheimia hemolytic
gb_Ba:AF314611	+	181.00	294.83	4.6e-08	6952	AF314611 Mannheimia hemolytic
gb_Ba:AF314613	+	181.00	294.83	4.6e-08	6952	AF314613 Mannheimia hemolytic
gb_Ba:AF314615	+	181.00	294.83	4.6e-08	6952	AF314615 Mannheimia hemolytic
gb_Ba:AF314617	+	181.00	294.83	4.6e-08	6952	AF314617 Mannheimia hemolytic
gb_Ba:AF314619	+	181.00	294.83	4.6e-08	6952	AF314619 Mannheimia hemolytic
gb_Ba:AF314621	+	181.00	294.83	4.6e-08	6952	AF314621 Mannheimia hemolytic
gb_Ba:AF314623	+	181.00	294.83	4.6e-08	6952	AF314623 Mannheimia hemolytic
gb_Ba:AF314625	+	181.00	294.83	4.6e-08	6952	AF314625 Mannheimia hemolytic
gb_Ba:AF314627	+	181.00	294.83	4.6e-08	6952	AF314627 Mannheimia hemolytic
gb_Ba:AF314629	+	181.00	294.83	4.6e-08	6952	AF314629 Mannheimia hemolytic
gb_Ba:AF314631	+	181.00	294.83	4.6e-08	6952	AF314631 Mannheimia hemolytic
gb_Ba:AF314633	+	181.00	294.83	4.6e-08	6952	AF314633 Mannheimia hemolytic
gb_Ba:AF314635	+	181.00	294.83	4.6e-08	6952	AF314635 Mannheimia hemolytic
gb_Ba:AF314637	+	181.00	294.83	4.6e-08	6952	AF314637 Mannheimia hemolytic
gb_Ba:AF314639	+	181.00	294.83	4.6e-08	6952	AF314639 Mannheimia hemolytic
gb_Ba:AF314641	+	181.00	294.83	4.6e-08	6952	AF314641 Mannheimia hemolytic
gb_Ba:AF314643	+	181.00	294.83	4.6e-08	6952	AF314643 Mannheimia hemolytic
gb_Ba:AF314645	+	181.00	294.83	4.6e-08	6952	AF314645 Mannheimia hemolytic
gb_Ba:AF314647	+	181.00	294.83	4.6e-08	6952	AF314647 Mannheimia hemolytic
gb_Ba:AF314649	+	181.00	294.83	4.6e-08	6952	AF314649 Mannheimia hemolytic
gb_Ba:AF314651	+	181.00	294.83	4.6e-08	6952	AF314651 Mannheimia hemolytic
gb_Ba:AF314653	+	181.00	294.83	4.6e-08	6952	AF314653 Mannheimia hemolytic
gb_Ba:AF314655	+	181.00	294.83	4.6e-08	6952	AF314655 Mannheimia hemolytic
gb_Ba:AF314657	+	181.00	294.83	4.6e-08	6952	AF314657 Mannheimia hemolytic
gb_Ba:AF314659	+	181.00	294.83	4.6e-08	6952	AF314659 Mannheimia hemolytic
gb_Ba:AF314661	+	181.00	294.83	4.6e-08	6952	AF314661 Mannheimia hemolytic
gb_Ba:AF314663	+	181.00	294.83	4.6e-08	6952	AF314663 Mannheimia hemolytic
gb_Ba:AF314665	+	181.00	294.83	4.6e-08	6952	AF314665 Mannheimia hemolytic
gb_Ba:AF314667	+	181.00	294.83	4.6e-08	6952	AF314667 Mannheimia hemolytic
gb_Ba:AF314669	+	181.00	294.83	4.6e-08	6952	AF314669 Mannheimia hemolytic
gb_Ba:AF314671	+	181.00	294.83	4.6e-08	6952	AF314671 Mannheimia hemolytic
gb_Ba:AF314673	+	181.00	294.83	4.6e-08	6952	AF314673 Mannheimia hemolytic
gb_Ba:AF314675	+	181.00	294.83	4.6e-08	6952	AF314675 Mannheimia hemolytic
gb_Ba:AF314677	+	181.00	294.83	4.6e-08	6952	AF314677 Mannheimia hemolytic
gb_Ba:AF314679	+	181.00	294.83	4.6e-08	6952	AF314679 Mannheimia hemolytic
gb_Ba:AF314681	+	181.00	294.83	4.6e-08	6952	AF314681 Mannheimia hemolytic
gb_Ba:AF314683	+	181.00	294.83	4.6e-08	6952	AF314683 Mannheimia hemolytic
gb_Ba:AF314685	+	181.00	294.83	4.6e-08	6952	AF314685 Mannheimia hemolytic
gb_Ba:AF314687	+	181.00	294.83	4.6e-08	6952	AF314687 Mannheimia hemolytic
gb_Ba:AF314689	+	181.00	294.83	4.6e-08	6952	AF314689 Mannheimia hemolytic
gb_Ba:AF314691	+	181.00	294.83	4.6e-08	6952	AF314691 Mannheimia hemolytic
gb_Ba:AF314693	+	181.00	294.83	4.6e-08	6952	AF314693 Mannheimia hemolytic
gb_Ba:AF314695	+	181.00	294.83	4.6e-08	6952	AF314695 Mannheimia hemolytic
gb_Ba:AF314697	+	181.00	294.83	4.6e-08	6952	AF314697 Mannheimia hemolytic
gb_Ba:AF314699	+	181.00	294.83	4.6e-08	6952	AF314699 Mannheimia hemolytic
gb_Ba:AF314701	+	181.00	294.83	4.6e-08	6952	AF314701 Mannheimia hemolytic
gb_Ba:AF314703	+	181.00	294.83	4.6e-08	6952	AF314703 Mannheimia hemolytic
gb_Ba:AF314705	+	181.00	294.83	4.6e-08	6952	AF314705 Mannheimia hemolytic
gb_Ba:AF314707	+	181.00	294.83	4.6e-08	6952	AF314707 Mannheimia hemolytic
gb_Ba:AF314709	+	181.00	294.83	4.6e-08	6952	AF314709 Mannheimia hemolytic
gb_Ba:AF314711	+	181.00	294.83	4.6e-08	6952	AF314711 Mannheimia hemolytic
gb_Ba:AF314713	+	181.00	294			

```

protein 20020 TR:AA54319 (EMBL:AE005178) (207 aa) fasta
scores: E(): 6.6, 26.3% id in 156 aa. The similarity is
limited to the central and C-terminal regions of this CDS"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/protein_id="CAC90808.1"
/db_xref="GI:15980009"
/gene
CDS
/translation="WAVVAKENTLITLNEPVSFLETSGTIEIKLRNYSISEVMSCV
AQLRLANKAHIFSIIIELEKTDHQSONYGAQGFVLYTPWTIRLMIIPGKAL
HOASAIPEYSEKVAHDADFIDLVGFGGYSYXEPNRIORTGLIQLSYKGEML
LYRENVETIYFKNKDVHIOHEPESLSVNLINKPDASAOQYFDEISDEGSLCQ
GKPHNSTIKTSORSITSLSTHGMNSKRAIEKIAINHNADEVRLSVYAMLRAAD
TVLSDVLTIEKSHYRVLAEINAAQGETSNAP"
complement(2167..3072)
/gene="yPO1996"
complement(2167..3072)
/gene="yPO1996"
/translation="No significant database hits"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/protein_id="CAC90809.1"
/db_xref="GI:15980010"
/translation="MSKVTSPDSIRIKHTEVIAASPNSESTISLSYKKEASEILE
KNAPFDLPLESLKRDVITIELNDLEKDTISNQILSEHVIYSRIEDNHLIM
RIGSPSTTLNANFEDILVNPMTPIPLQYQVCEDESPRSLIRIEDNPLMP
GVYVEFAKYLIDFETAAQDAFLIHASPKGLSWYKNSLEPSTICTSLAS
RLOMYVLMGMANTQAVPLEKLAKSYTANFVREAVESLIDPQSTLELDLYVK
NNSDAQIIEAASQALIKAESEV"
complement(3863..4936)
/gene="yPO1997"
complement(3863..4936)
/gene="yPO1997"
/translation="No significant database matches"
/codon_start=1
/translation="putative exported protein"
/product="putative exported protein"
/protein_id="CAC90810.1"
/db_xref="GI:15980011"
/translation="MKALKITTCMLVGLISOSTAEPTKPLVIOSGSFAAGAV
ITTPROFPRKPIDPTGOTYHGDHASVYQIPEPHKYPITMLGAQSSSTWSTEDP
GREGRONTILRKRGSTYLLVDOPRRGNAGRSNVEGTIYKLADEQLMFNQFRIQWPDFW
DSVQSHDEETLNQYRYQMTPTGPFQIDVLSDAASAVDSGPAILLFTSHOGGPGPM
FPAKNANANVGIYAFEPSSFEVPEGDVPPPTKNAFDVKGESVPLEKFMAITRIPIVLA
IYGCINIPKPKTPALPAODSMRRLAVARQMRDVYVNHCGDVTYIHLPEIGIKGNTHPP
FSDLNVEETADVSKEFLKDNLO"
complement(5149..5619)
/gene="yPO1998"
complement(5149..5619)
/gene="yPO1998"
/translation="Similar to Escherichia coli Pep1 TnpA TR:p75026
(EMBL:U60777) (140 aa) fasta scores: E(): 0.13, 31.4% id
in 137 aa, and to Thermotoga maritima hypothetical 15.1
kDa protein Tm1010 TR:Q9X0A3 (EMBL:AB001162) (135 aa)
fasta scores: E(): 1.2e-12, 42.0% id in 119 aa"
/codon_start=1
/translation="putative exported protein"
/product="putative exported protein"
/protein_id="CAC90811.1"
/db_xref="GI:15980012"
/translation="MKVFLITLALLLPGTQALAAQAAYVSPAGSQPTTIGAPESF
TGTAVLDSFKGTAPARISGTVSEAGAPFAMTHPIPGOTLIYSSGTGWQMEGAA
QOIKGADVWIPPGYKWHMGASAKSESVHIAVSESLDQNTVTMKEKVTDDGYPq"
complement(5674..6096)
/gene="yPO1999"
complement(5674..6096)
/gene="yPO1999"
/EC_number="4.1.1.44"
/translation="Similar to Acinetobacter calcoaceticus
4-carboxymuconolactone decarboxylase Pcad SW:DD4C.ACT1A
(P20370) (134 aa) fasta scores: E(): 1.4e-09, 35.9% id in
103 aa, and to Methanobacterium thermophilicum

```

```

gamma-carboxymuconolactone decarboxylase Mch234 Tr:026333
(EMBL:AE000810) (123 aa) fasta scores: E(): 4.7e-11, 35.5%
id in 107 aa"
/codon_start=1
/transl_table=11
/product="putative decarboxylase"
/protein_id="CAC90812.1"
/db_xref="GI:15980013"
/translacion="MMKKKIQYLLCTALTTPAGSMASIAHAODAKNNKTPSPHAQDL
MGDIAPMAQDLTDVLADWIERPOLSRDRSLVTYSALVAMPDRDRLRSHLRARON
GVPEELVEALITOLAFYSGMPTAVTATLSAKEYKEKG"
complement(6373. .7764)
/gene="YP02000"
complement(6373. .7764)
/gene="YP02000"
/EC_number="2.7.3.-"
/note="Similar to Burkholderia pseudomallei sensor protein
Irls SW:IRLS.BURPS (031396) (464 aa) fasta scores: E(): 0,
29.1% id in 454 aa, and to Pseudomonas aeruginosa probable
two-component sensor PAL438 Tr:AG604827 (EMBL:AE004573)
(481 aa) fasta scores: E(): 0, 42.2% id in 453 aa"
/codon_start=1
/codon_end=1
/transl_table=11
/product="putative two-component system sensor protein"
/protein_id="CAC90813.1"
/db_xref="GI:15980014"
/translacion="MMIKRSISGRILAMFALISILVSALGIVLRSSLNHSLOKMH
NELOFTESIMAPWILSRITLDDGMOTLAKTFTDLAAAGSGVQWVLSDDPOLQVGPFA
PRGVMTASLKGKSPKVGAGNDESCSLFMSLEIAPGERRALRYVVAIDSTPYMGNTLN
EFTFTLVITATIGVLPFALLAGFYATRGMPYKLSQAOAHNLARGNNGQALDITLPLD
ELQALVAENGVLSEKQETIAWRQLESFNADVAHNEIRPLTILMCOQTQGLSKRKEHL
EDLESNLEELERMTSIVNDMLFSAQAGEHATQLDVSLREYTKTTEYVERSEFAE
KNLSIHVEGDITATIDRRFLRSFLANLLENSARHAYGVTVKVTISNNMLASIAVAN
IGAPIAEHLNRLPERFRRVDSNRASDTHHGLSLIVRAVALMHNRDVPASSQNGIN
TPGTLTITDPYI"
complement(6334. .7017)
/gene="YP02000"
/note="Pfam match to entry PF00512 signal, Histidine
kinase, score 129.50, E-value 6.1e-35"
complement(7057. .7269)
/gene="YP02000"
/note="Pfam match to entry PF00672 DUF5, HAMP domain,
score 52.70, E-value 7.8e-12"
complement(7213. .7278)
/gene="YP02000"
/note="1 probable transmembrane helices predicted for
YP02000 by TMHMM2.0"
complement(7770. .8462)
/gene="copr"
/note="copr"
/note="YP02001"
complement(7770. .8462)
/gene="copr"
/note="copr"
/note="Similar to Pseudomonas syringae transcriptional
activator protein COPR SW:COPR.PSESM (002540) (227 aa)
fasta scores: E(): 0, 60.8% id in 227 aa, and to
Burkholderia pseudomallei transcriptional activator
protein Irlr SW:IRLR.BURPS (031395) (229 aa) fasta scores:
E(): 0, 57.6% id in 224 aa"
/codon_start=1
/transl_table=11
/product="transcriptional activator"
/protein_id="CAC90814.1"
/db_xref="GI:15980015"

```



```

CDS
3867. .4112
/gene="XF2399"
/feature="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AF85198.1"
/db_xref="GI:9107582"
/translation="MPINTTASHKTHRPCCOHCNDASATATATPTTPTPTPL
KDSRAPLAIATACLRPPKTHHAIHVPYITTK"
4428. .5057
/gene="XF2400"
4428. .5057
/gene="XF2400"
/feature="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AF85199.1"
/db_xref="GI:9107583"
/translation="MTMTANVFKKILKICGMVIALMAFGAFISGIIYIKKAO
DATAKIKESNETILMIGSDSSYKRCFKDELATGLGIIIFNNELYSVEK
SPKPHGPPSSVTIGIFGKYSKKNYINSYNQDONKQITVIGIKQELLYEFSNT
YSAMDNDLFEVIEYKADGQPKKKERHKLIDLEAHITPCOGES"
5054. .5326
/gene="XF2401"
5054. .5326
/gene="XF2401"
/feature="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AF85200.1"
/db_xref="GI:9107584"
/translation="NMTSSYFFEDATLDRAEYRSGNCPGMYQILSDAVKAGGDA
RLANGLNAGNDHISAQGRDPSIRAMDADREKNINTQHHPAL"
5329. .5946
/gene="XF2402"
5329. .5946
/gene="XF2402"
/feature="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AF85201.1"
/db_xref="GI:9107585"
/translation="MNSPRVFKSLKLCGLALTVLMVFGAILSCGNMTEETTPORA
AAODIKESKMMLOSRNSKRCFEPVWPVOPSKAIFPKNIYLSIETSP
GNLGEAGTSIVLIPDPKSKKEYIHAIPLGNSNOVEHGYKTFYNNLNDPLXS
AIDKGTLENIYEQOPRNKKEWTHDLIDFVNHTIPCOGES"
5943. .6215
/gene="XF2403"
5943. .6215
/gene="XF2403"
/feature="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AF85202.1"
/db_xref="GI:9107586"
/translation="MWTSLDFYDQATLERAMTYNNGDMAMYQTLSDAVNAGGDA
RLANGFKNAGDIHSPQRDTPSIRAMDADREKNINTQHHPAL"
6218. .6841
/gene="XF2404"
6218. .6841
/gene="XF2404"
/feature="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"

```

```

/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AF85203.1"
/db_xref="GI:9107587"
/translation="MNSPRVFKSLKLCGLALTVLMVFGAILSCGNMTEETTPORA
AAODIKESKMMLOSRNSKRCFEPVWPVOPSKAIFPKNIYLSIETSP
GNLGEAGTSIVLIPDPKSKKEYIHAIPLGNSNOVEHGYKTFYNNLNDPLXS
AIDKGTLENIYEQOPRNKKEWTHDLIDFVNHTIPCOGES"
6838. .7110
/gene="XF2405"
6838. .7110
/gene="XF2405"
CDS
alignment_scores:
Quality: 241.00 Length: 98
Ratio: 3.347 Gaps: 0
Percent similarity: 73.469 Percent identity: 53.061
alignment_block:
US-09-484-577A-4 x AE004049
Align seg 1/1 to: AE004049 from: 1 to: 15552
20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2495 CGCAGCGCGATGAGCGACCTCTCCGCCCATCTAGAACTGACCGA 2544
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAla 53
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2545 AACCCCTGTCCTCCCTACCGCGCGATGCGCATCTATTATTCCT 2594
53 eupThrCysAlaValAlaArgPalaGlyLeuGlyArgIleAspLeuG1 69
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2555 GCTTATGCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2644
70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2645 GCGGTCGCGCGCGGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2694
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyG1 103
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2695 ACCACTAGACACCGCGTAGTACCGCTATCTGCGGCAAGTACGCA 2744
103 hrValLysAlaGlyLurLeuIleGluLeuAspProPheAla 117
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2745 AAGTGAAGACACACCACTGATGCAACTGATGCCACCGCC 2788
seq_name: gb_ba:NM391263
seq_documentation_block:
LOCUS NM391263 1428 bp DNA linear BCT 29-MAR-2000
DEFINITION Neisseria meningitidis hlyd gene (putative component of type I
secretion system), strain FAM18.
ACCESSION AJ391263 GI:6900444
VERSION AJ391263.1
KEYWORDS hlyd gene.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 1428)
AUTHORS Klee,S.R., Nassif,X., Kusecek,B., Merker,P., Beretti,J.L.,
Achtman,M. and Tinsley,C.R.
TITLE Molecular and biological analysis of eight genetic islands that
distinguish neisseria meningitidis from the closely related
pathogen neisseria gonorrhoeae
JOURNAL Infect. Immun. 68 (4), 2082-2095 (2000)
MEDLINE 20187481
REFERENCE 2 (bases 1 to 1428)
AUTHORS Achtman,M.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-2000) Achtman M., Abt. Trautner,

```

Max-Planck-Institut fuer molekulare Genetik, Innestr. 73, Berlin,  
14195, GERMANY

# FEATURES

Location/Qualifiers

1..1428

/organism="Neisseria meningitidis"

/strain="FAM18"

/db\_xref="taxon:487"

/country="USA"

/note="serogroup C, Et-37 complex"

1..1428

/gene="hlyD"

1..1428

/function="putative inner membrane component of the type I secretion system"

/note="99% identity to NacD protein of N. meningitidis serogroup C strain IR1075; homology to hemolysin secretion protein P09986"

/codon\_start=1

/transl\_table=11

/product="HlyD protein"

/protein\_id="CA872043.1"

/db\_xref="GI:6900445"

/translation="MFFSALKSFSLRYITVWRNVAVRDQLEPPKRTAEQALPAHL  
ELTDFPVSAPKMAARFIMAFALLALIMSWFKIDIVAAASGTVSGRSKTIOPLET  
AVKAVHVRDGOHVKOGETLAEVGTDSVDVSEQAOLAOOLSKRYEAVLALES  
RVPHIDMAOARSLGSDADVOAOLAOHQAQVAAOAOLOALRHOAELOSAKA  
OEKLVSVGAIEQKTAADYRRLRANFISPHAFLEQSKSVNMNDESTRGOKOTO  
AAIAQESRVLNTONLKRDTDLAKOANQIDQYRQDTKAKORQOOLMTIQSPADGT  
VOELATYVGVVQAQAKMVAAPDDDKMDEVLYLNKDIGVEQODAVVKTISFPY  
TRGYLTGVKSVSHDASHQELGLVYAVVSLDKHTLIDGRAVNLTAQMYTAETK  
TGKRRVLDLSPLOTKIDSFRR"

BASE COUNT 338 a 322 c 473 g 295 t

ORIGIN

alignment\_scores:

Quality: 238.00 Length: 103

Ratio: 3.216 Gaps: 0

Percent Similarity: 71.845 Percent Identity: 49.515

alignment\_block:

US-09-484-577A-4 x NME391263 ..

Align seg 1/1 to: NME391263 from: 1 to: 1428

20 LysTharGaspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36

CGACGCGCGGAAAGACAGGCGTTTGGCGGCAATTTGGAAGTACCGG 143

36 uThrProProSerProThrAlaArgLeuThrAlaAlaLeuAlaAla 53

144 TACGCGGCTCTGCGCGTCCGAAATGGCGCGCTTTATATATATGCGGT 193

53 euphetyCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69

194 TTGGCGCTTTGGCTTTGTTGTCCTGCTGCGGAAATCGATATATGTC 243

70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86

244 GCGGCGGCTTGGCGGAAAGCGTGTGCGGCGCGGCGGCGGCGGCGGCGG 293

86 nProLeuGluValGlyValAlaValArgAlaThrHisValArgAspGlyGlnT 103

294 GCGCGTGGAAAGCGCGGTGTTAAGCGGTACATGTGCGGCGGCGGCGGCGG 343

103 hrValLysAlaGlyIleLeuIleGluLeuAspProPheAlaGlyGly 119

344 ATGTGAACAGGAGAAACGTCGCGAAGTGAAGCTGTGGAACAGAC 393

120 ValAspVal 122

394 AGCGATGTG 402

seq\_name: gb\_AFI21772

seq\_documentation\_block:

LOCUS AFI21772 4254 bp DNA linear BCT 18-MAY-1999

DEFINITION Neisseria meningitidis NatD (natD) and NatC (natC) genes, complete cds.

ACCESSION AFI21772

VERSION AFI21772.1 GI:4838368

KEYWORDS

SOURCE

ORGANISM

Neisseria meningitidis.

Neisseria meningitidis

Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

Neisseria.

1 (bases 1 to 4254)

Stojiljkovic,I.

Neisseria meningitidis-specific gene cluster encoding components of type I secretion system

Unpublished

2 (bases 1 to 4254)

Stojiljkovic,I.

Direct Submission

Submitted (21-JAN-1999) Microbiology and Immunology, Emory School of Medicine, 1510 Clifton Rd., Atlanta, GA 30322, USA

Location/Qualifiers

1..4254

/organism="Neisseria meningitidis"

/strain="IR1075"

/db\_xref="taxon:487"

/map="3' of rfa gene and 5' of the putative rfa homolog"

/note="serogroup: C"

218..1645

/gene="natD"

218..1645

/gene="natD"

/note="component of the type I secretion system; Actinobacillus actinomycetemcomitans leukotoxin secretion protein D (HlyD) homolog"

/codon\_start=1

/transl\_table=11

/product="NatD"

/protein\_id="AA03097.1"

/db\_xref="GI:4838369"

/translation="MFFSALKSFSLRYITVWRNVAVRDQLEPPKRTAEQALPAHL  
ELTDFPVSAPKMAARFIMAFALLALIMSWFKIDIVAAASGTVSGRSKTIOPLET  
AVKAVHVRDGOHVKOGETLAEVGTDSVDVSEQAOLAOOLSKRYEAVLALES  
RVPHIDMAOARSLGSDADVOAOLAOHQAQVAAOAOLOALRHOAELOSAKA  
OEKLVSVGAIEQKTAADYRRLRANFISPHAFLEQSKSVNMNDESTRGOKOTO  
AAIAQESRVLNTONLKRDTDLAKOANQIDQYRQDTKAKORQOOLMTIQSPADGT  
VOELATYVGVVQAQAKMVAAPDDDKMDEVLYLNKDIGVEQODAVVKTISFPY  
TRGYLTGVKSVSHDASHQELGLVYAVVSLDKHTLIDGRAVNLTAQMYTAETK  
TGKRRVLDLSPLOTKIDSFRR"

1714..3117

/gene="natC"

1714..3117

/gene="natC"

/note="component of the type I secretion system; Serratia marcescens HasF, Pseudomonas aeruginosa AprF and Escherichia coli TolC protein homolog"

/codon\_start=1

/transl\_table=11

/product="NatC"

/protein\_id="AA030978.1"

/db\_xref="GI:4838370"

/translation="MTLLNLMIMQDYGISVCLTLPYLQHELSAMKSYKYLTPS  
LFTPLISPSVSAFTLPEAWRAAOHSDAFQASHVQRODAVRAAOQAKAFLPHVSA  
MSYQROPBSISSTRETQGSVOVGOTLPAARFAQVROSRFPOQAQAFORDAREEL  
LKVASSTFENVLSRDTVAHAHAKEKAYAQVQAOALPKGAATFALDIHAKAGYN  
ALAOELAVLAQKOTYEMQNDYDLDKQLEADTANLALRYLTKLRSYLDENQRTA  
LSNNHEYRMQOLIOSGQALRAAONSRYPTVSAHVQYNNLYTSSAONNDYHYRGGS  
MSYVQNLPLETGGELSGKIHAEAOYGAALQVLTATERRHILKAVQAATSEGAARY  
QIMAOERVLSSRLKSTETGGQYGIARNRLVIRARQVQAOBOKIQAQRYKFMLAY  
LRLVKSGLGLTVFAE"

```

source
    3200..4254
    /note="insertion sequence; similar to the putative
    /note="synchocystis PCC6803 transposase and small repeated
    /note="elements of Neisseria gonorrhoeae"
    /organism="Neisseria meningitidis"
    /db_xref="taxon:487"

BASE COUNT      1070 a      1021 c      1130 g      1032 t      1 others
ORIGIN

alignment_scores:
    Quality:      238.00      Length:      103
    Ratio:         3.216      Gaps:         0
    Percent Similarity: 71.845      Percent Identity: 49.515

alignment_block:
    US-09-484-577A-4 x AF121772 ..

Align seg 1/1 to: AF121772 from: 1 to: 4254

20  LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuIleValG1 36
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
311 CGCAGCGCGGAGAACAGCGCTTTTCCCGCGCATTTGGAACTGACCGCA 360

36  uThrProProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
361 TACGGCGGCTCTGCGCTCCGAAATGGCGCGGCTTTATTATGCGCT 410

53  euPheTyrCysAlaValAlaTrpAlaGlyLeuGlyAlaGlyIleVal 69
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
411 TTGGCGCTTTGGCTTGTGTTGCTGCTGCTGCGCAATGCAATGATG 460

70  AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
461 GCGGCGGCTGCGGCAAAACGCTGCGGCGGCGGCGGCAAAACCATCCA 510

86  nProLeuGluAlaGlyValAlaArgAlaThrHisValArgAspGlyInt 103
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
511 GCGCGTGAACAGCGCGGTGTTAAAGCGGTACATGTGCGCGGCGGCGAGC 560

103 hrValLysAlaGlyIleLeuIleLeuIleGluLeuAspProPheAlaGly 119
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
561 ATGTGAACAGCGGAGAACGCTGCGGCAACTGAGCGCTGTGGACACAGAC 610

120 ValAspVal 122
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
611 AGCGATGTG 619

name: gb_ba:AE002524

seq documentation block:
LOCUS      AE002524      11381 bp      DNA      linear      BCT 25-MAY-2000
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 166 of 206
ACCESSION  AE002524 AE002098
VERSION    AE002524.1 GI:7226991
KEYWORDS
SOURCE
ORGANISM   Neisseria meningitidis MC58.
            Neisseria meningitidis MC58
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE  1 (bases 1 to 11381)
AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Clitane,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Yamathavan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizzi,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
            Complete genome sequence of Neisseria meningitidis serogroup B
            strain MC58

JOURNAL    Science 287 (5459), 1809-1815 (2000)
MEDLINE    2017575
PUBMED     10710307
REFERENCE  2 (bases 1 to 11381)
AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Clitane,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Yamathavan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizzi,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
            Direct Submission
            Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
            Location/Qualifiers
                1..11381
                /organism="Neisseria meningitidis MC58"
                /strain="MC58"
                /db_xref="taxon:122586"
                /note="serogroup: B"
                165..976
                /gene="NMB1736"
                /note="This region contains an authentic point mutation,
                causing a premature stop, and is not the result of a
                sequencing artifact. This region contains an authentic
                frame shift and is not the result of a sequencing
                artifact; similar to PID:1653908 PID:1651692"
                complement(1140..2543)
                /gene="NMB1737"
                complement(1140..2543)
                /gene="NMB1737"
                /note="similar to PID:1139569 percent identity: 48.26;
                identified by sequence similarity; putative"
                /transl_table=1
                /product="secretion protein, putative"
                /protein_id="AAF42081.1"
                /db_xref="GI:7226992"
                /translation="MTLNLMTIMODYGISVCTLTPTLQHEFSAKMSYFSKYILPVS
                LFTPLSLSPSVSAFTLPEAMRAAQHSADFOASHQROAVRARRQQAAPLPHVSA
                NASYQSPYNSISSTRETQGSVYVQGLTFDAFAQYRQSFQTAQAEADAREEL
                LKVAESYFNVLSDRTVAHAHAKEAYAOVQAOLFNKGAATLADIHEAAGADN
                ALAOLFIAVLAERQYENOLNDYTDLSKQIEATDITLARIYKTERYSIDMORITA
                LSNHEYRMQOLALOSSGQALRAQNSRYPAHYAHGYQNNLTSSAONNDYTRKRG
                MSYVOLNLPITYGEGELSGKIEHAERQYGAERQQLATERHITLAVRQAYTESGAR
                YQIMQERVLSSRLKLTSTETGOYGINRNLLEYIRAREVAQAQKLAQRYKFMILAY
                LRLVKSGLGLTFVFAE"
                complement(2612..4039)
                /gene="NMB1738"
                complement(2612..4039)
                /gene="NMB1738"
                /note="similar to GP:558154 percent identity: 65.47;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="secretion protein, putative"
                /protein_id="AAF42082.1"
                /db_xref="GI:7226993"
                /translation="MFESALKSFSLRYITVWNVAVRDQLPKPKRTAEQAFPLAHL
                ELTDPVPSAPKWAAREITMAFALLALMLSMFQRIDIVAAASQTVSGGRSKTQPLET
                AVYKAVHVRDQGHVQGETLALEAVLQSDVYOSQALQALQSLREAVLAALES
                RTVPHIDMAQARSILGSDADVQSAOVLQHOVQANVAQOALQSLRGAQELQSKA
                QOQKLVSAITQOQTADYRRARRANFSEHAFLEQOQSVSNMNDLESTROMRQIO
                AALAEQONRYVNTQNTKQTDIDALRQANEQIDQTRGQTDKAKQROQLTQSPADGT
                VOELATITGVGVYAQAQMMVIAPDQDKMDVAVLVNMDIGFVEGQDVAVVTIESPY
                TRYGLTGVKSVSDAASHQQLGVYAVVSLDKHTNLINDKAVNLTAGMVAETAIK
                TGRRLVLYLTLSPQTKLDESFRER"
                4275..4628
                /gene="NMB1739"
                4275..4628
                /gene="NMB1739"

```

```

/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42083.1"
/db_xref="GI:7226994"
/translation="MTTNDLRHFLGGLAIFSGYWGTMILLMFLVREAYKPKRN
BKILSGSITAAAILMLFWIITPKQPIKEELQAGBENDRKKEAAVANEQCKTAG
BETRRIRIMWKGLCC"
4752..5162
/gene="NMB1740"
4752..5162
/gene="NMB1740"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42084.1"
/db_xref="GI:7226995"
/translation="MGNTAMMSNCNPTPIPLGIGQKIPPLNQTFFHHPARVATPE
NNVDSKLRHWWAGATRIIDQTDVIAKTIYVEKGLDGTGAGMPKFAILCNK
BRLTSEPLSDVLSYLKPYILRLPYIASLRDD"
5233..6575
/gene="NMB1741"
/note="this region contains an authentic frame shift and
is not the result of a sequencing artifact; conserved
hypothetical protein, authentic frameshift; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
6620..6910
/gene="NMB1742"
6620..6910
/gene="NMB1742"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42085.1"
/db_xref="GI:7226996"
/translation="NNILOGHIVGVKINGQIEGKSPDYCCLIIVATPDSQGNALG
SSTPEVDGGSANFEQFRNAQFLEANLVEIYTTGKTKLVIGFQLVKKG"
6915..7112
/gene="NMB1743"
6915..7112
/gene="NMB1743"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42086.1"
/db_xref="GI:7226997"
/translation="MOKYVVOVSSTGDFLYLSPETGDIHTKLTINADYFDEEAI
NGLLEIGNQYEEVVEFLKD"
7120..7404
/gene="NMB1744"
7120..7404
/gene="NMB1744"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42087.1"
/db_xref="GI:7226998"
/translation="MFGGRKNAPSITAKHPLKFNIMKFINTCRKYGAKLAVTAAPL
ALAAHANATLPDRAKNLLEAKADGMEAGWLVGIFALFVFSIVKRYMK"
7411..7689
/gene="NMB1745"
7411..7689

```

```

/gene="NMB1745"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42088.1"
/db_xref="GI:7226999"
/translation="MYOVGNCKLEKHOAENLYSLVPRIKENQIVREYNGSLMK
MSDQPIRILLAECSKPNQNLQSGLETGMIVFGILASVYVSLKKYK"
7817..8107
/gene="NMB1746"
7817..8107
/gene="NMB1746"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42089.1"
/db_xref="GI:7227000"
/translation="NNKNRDRFFYISNSDLKLSGSPDRPLSYRYCYLKEFGLLKN
FSMDKCHFFNRIINFNECSCEIKFDDSFITGNKIDVSDSNRFSVSPEC"
8148..9653
/gene="NMB1747"
8148..9653
/gene="NMB1747"
/note="similar to GP:3758871 percent identity: 86.60;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="tsbp protein, putative"

```

alignment\_scores:

Quality:	238.00	Length:	103
Ratio:	3.216	Gaps:	0
Percent Similarity:	71.845	Percent Identity:	49.515

alignment block:

US-09-484-577A-4 x AE002524/rev ..

Align seg 1/1 to reverse of: AE002524 from: 1 to: 11381

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaLeuGluIleValG1 36
:::||||| ::||| ||||| ||||| ||||| :::::
3946 CGCAGCGCGGAGAGACAGCGCTTTTGCGCGCATTTGGAACAGCGA 3897
36 uThProProSerProThralaArgLeuThAlaAlaLeuAlaIal 53
::||| ||| ::::: ||| ::::: |||
3896 TACGCGGCTCTGCTGCCGAATGCGCGCGGCTTTATATGCGGT 3847
53 eUpheTyrcysAlaValAlaThrPalaGlyLeuGlyIleAspIleVal 69
::||| ||| ::||| ::||| ::||| ::|||
3846 TTGCGCTTTGGCTTTGTGTGTGCTGCTGCGGCAAAATCATATTGTG 3797
70 AlaseAlaseArgLysIleValProGlyAspArgValLysLeuValG1 86
|||::||| ||| ||| ||| ||| ||| |||
3796 GCGCGGCTTTGCGGCAAAACGCTGTGCGCGCGGCGGCAAAACATCA 3747
86 nProLeuGluValGlyValAlaArgAlaThrHisValaArgAspGlyInt 103
||||| ::||| ::||| ::||| ::||| ::|||
3746 GCCCGTGGAAACGCGGCGGTAAAGCGGTACATGTGCGCGCGGCGG 3697
103 hValLysAlaGlyGluLeuIleGluLeuAspProPheAlaGlyGly 119
||||| ||| ||| ||| ||| |||
3696 ATGTGAACACAGGAGAAACGCTGGCGGAACGTGTGGAACAGAC 3647
120 ValAspVal 122
|||||
3646 AGCGATGTG 3638
seq_name: gb_pat:AX044033

```



seq\_documentation\_block: 349980 bp DNA linear PAT 24-NOV-2000  
LOCUS AX044033  
DEFINITION Sequence 112 from Patent WO0066791.  
ACCESSION AX044033  
VERSION AX044033.1 GI:11342917  
KEYWORDS  
SOURCE Neisseria meningitidis.  
ORGANISM Neisseria meningitidis.  
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
Neisseria.  
REFERENCE 1 (bases 1 to 349980)  
AUTHORS Piazza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,  
Maignani, V., Galeotti, C., Mora, M., Ratti, G., Scariselli, M.,  
Scarlatto, V., Rappuoli, R., Frazer, C. M. and Grandi, G.  
TITLE Neisseria genomic sequences and methods of their use  
JOURNAL Patent: WO 0066791-A 112 09-NOV-2000;  
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)  
FEATURES  
Source Location/Qualifiers  
1..349980  
/organism="Neisseria meningitidis"  
/db\_xref="taxon:487"  
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to  
349980 349980 bases-seq 108: 300001 to 649980 349980  
bases-seq 109: 600001 to 949980 349980 bases-seq 110:  
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980  
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq  
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to  
2272325 172325 bases"  
BASE COUNT 87189 a 93501 c 84627 g 84663 t  
ORIGIN  
alignment\_scores: Quality: 238.00 Length: 103  
Ratio: 3.216 Gaps: 0  
Percent Similarity: 71.845 Percent Identity: 49.515  
alignment\_block:  
US-09-484-577A-4 x AX044033/rev ..  
Align seg 1/1 to reverse of: AX044033 from: 1 to: 349980  
20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36  
:::||||| ::||| ||||| ||||| ||||| :::::  
323829 CGCAGCGGGGAGAAACAGCGCTTTTGGCCCGCATTTGGAACGACCGA 323780  
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53  
::||| ||| :::: ||| ::||| ::||| ::||| ::|||  
23779 TACGCGGCTCTGCGCTCGGAATGCGCGCTTTTATATGCGCT 323730  
53 eupHeTyCyAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69  
::||| ||| ::||| ::||| ::||| ::||| ::||| ::|||  
323729 TTGGCGCTTTGGCTTTGTTGGTCGCTCGCAAAATCGAATATGTCG 323680  
70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86  
|||:::||||| ||| ||| ||| ||| ||| ::|||  
323679 GCGGCGGCTTTGGGCAAAACGCTGCGGCGGCGGCAACCAACATCCA 323630  
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyG1NT 103  
|||:::||||| ::||| ::||| ::||| ::||| ::||| ::|||  
323629 GCCCTGGAACAGCGCGGTGTTAAGCGGTACATGTGCGGCAACGCGCAGC 323580  
103 hrValLysAlaGlyLueIleuIleGluLeuAspProPheAlaGlyGly 119  
||| ||| ||| ||| ||| ||| ||| ||| ::|||  
323579 ATGTGAACAGAGGAAACGCTGCGGCAATGAGAGCTGTGGAAACAGAC 323530  
120 ValAspVal 122  
||| |||  
323529 AGCGATGTG 323521  
seq\_name: gb\_pat:AX044034  
seq\_documentation\_block:

LOCUS AX044034 349980 bp DNA linear PAT 24-NOV-2000  
DEFINITION Sequence 113 from Patent WO0066791.  
ACCESSION AX044034  
VERSION AX044034.1 GI:11342918  
KEYWORDS  
SOURCE Neisseria meningitidis.  
ORGANISM Neisseria meningitidis.  
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;  
Neisseria.  
REFERENCE 1 (bases 1 to 349980)  
AUTHORS Piazza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,  
Maignani, V., Galeotti, C., Mora, M., Ratti, G., Scariselli, M.,  
Scarlatto, V., Rappuoli, R., Frazer, C. M. and Grandi, G.  
TITLE Neisseria genomic sequences and methods of their use  
JOURNAL Patent: WO 0066791-A 113 09-NOV-2000;  
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)  
FEATURES  
Source Location/Qualifiers  
1..349980  
/organism="Neisseria meningitidis"  
/db\_xref="taxon:487"  
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to  
349980 349980 bases-seq 108: 300001 to 649980 349980  
bases-seq 109: 600001 to 949980 349980 bases-seq 110:  
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980  
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq  
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to  
2272325 172325 bases"  
BASE COUNT 86473 a 95646 c 85908 g 81953 t  
ORIGIN  
alignment\_scores: Quality: 238.00 Length: 103  
Ratio: 3.216 Gaps: 0  
Percent Similarity: 71.845 Percent Identity: 49.515  
alignment\_block:  
US-09-484-577A-4 x AX044034/rev ..  
Align seg 1/1 to reverse of: AX044034 from: 1 to: 349980  
20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36  
:::||||| ::||| ||||| ||||| ||||| :::::  
23829 CGCAGCGGGGAGAAACAGCGCTTTTGGCCCGCATTTGGAACGACCGA 23780  
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53  
::||| ||| :::: ||| ::||| ::||| ::||| ::|||  
23779 TACGCGGCTCTGCGCTCGGAATGCGCGCTTTTATATGCGCT 23730  
53 eupHeTyCyAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69  
::||| ||| ::||| ::||| ::||| ::||| ::||| ::|||  
23729 TTGGCGCTTTGGCTTTGTTGGTCGCTCGCAAAATCGAATATGTCG 23680  
70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValG1 86  
|||:::||||| ||| ||| ||| ||| ||| ::|||  
23679 GCGGCGGCTTTGGGCAAAACGCTGCGGCGGCGGCAACCAACATCCA 23630  
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyG1NT 103  
|||:::||||| ::||| ::||| ::||| ::||| ::||| ::|||  
23629 GCCCTGGAACAGCGCGGTGTTAAGCGGTACATGTGCGGCAACGCGCAGC 23580  
103 hrValLysAlaGlyLueIleuIleGluLeuAspProPheAlaGlyGly 119  
||| ||| ||| ||| ||| ||| ||| ||| ::|||  
23579 ATGTGAACAGAGGAAACGCTGCGGCAATGAGAGCTGTGGAAACAGAC 23530  
120 ValAspVal 122  
||| |||  
23529 AGCGATGTG 23521  
seq\_name: gb\_da:NME391260  
seq\_documentation\_block:  
LOCUS NME391260 7824 bp DNA linear BCT 29-MAR-2000

```

DEFINITION
Neisseria meningitidis DNA for region 7 (homology to type I
secretion system) and flanking genes, strain 22491.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1 (bases 1 to 7824)
Klee,S.R., Nassif,X., Kusecek,B., Merker,P., Beretti,J.L.,
Achtman,M. and Tinsley,C.R.
Molecular and biological analysis of eight genetic islands that
distinguish neisseria meningitidis from the closely related
pathogen neisseria gonorrhoeae
Infect. Immun. 68 (4), 2082-2095 (2000)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
20187481
2 (bases 1 to 7824)
Achtman,M.
Direct Submision
Submitted (19-JAN-2000)
Max-Planck-Institut fuer molekulare Genetik, Ihnestr. 73, Berlin,
14195, GERMANY

FEATURES
source
1..7824
Location/Qualifiers
/organism="Neisseria meningitidis"
/strain="22491"
/db_xref="taxon:487"
/country="Gambia"
/note="serogroup A, subgroup IV-1"
complement(1..1002)
/note="orf-C7, homology to AB015669; designated NM1999 in
the Sanger Center 22491 genome sequence"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAB72024.1"
/db_xref="GI:6900415"
/translation="MLAEVVGFGIYKESKGGGRFYERCMVLMGSEDLVYGRVHFGGO
NNILFELTGTGCGVAKRGWESRLFAFLTNAIRKIRVDIAXDPFNGEYSPNARD
RNKMFCTCHHWKPKGEGCSIDWEDDAKMKRGTYGIGSRSSKYYRVYKGGKOLD
KTSYWTREIEFFKAKDIYIPREVONNGEYRGAYPICERVOKATRIHAKEDKVIS
ADRLEWYKKGFGAANKIKFIPELDIAKIFELIESSHKPLSLAPAVDCAFLKA
QAIHQEPAKFKYKDPYIMYETYLENLEKQLEQKKYNNDESINNTYDKFARPLSMA"
1456..6701
/note="Region 7, a DNA island present in N. meningitidis
and absent from N. gonorrhoeae"
2510..3874
/gene="hlyD (left fragment)"
2510..3874
/gene="hlyD (left fragment)"
/gene="hlyD (left fragment)"
/note="homology to hemolysin secretion protein D P09986;
98% identical to NatD component of the type I secretion
system AAD30977; designated natD' (NM1996) in the Sanger
Center genome sequence"
/codon_start=1
/pseudo
/transl_table=1
3820..3953
/note="insertion/duplication that inactivates hlyD gene"
3907..4071
/gene="hlyD (right fragment)"
3907..4071
/gene="hlyD (right fragment)"
/gene="hlyD (right fragment)"
/note="designated natD' (NM1995) in the Sanger Center 22491
genome sequence"
/codon_start=1
/pseudo
/transl_table=1
4140..5543
/gene="tolC"
4140..5543

```

[illegible]

```

2603 CGCAGCGGCGGAGACAGCGCTTTTGGCCGCCCATTTGGAACTGACCGA 2652
36 uThrProSerProThrAlaArgLeuThrAlaLeuAlaLeuAlaAla 53
2653 TAGCGCGGCTCTGCCGCTCCGAAATGGCGCGCTTTTATTATGAGCGT 2702
53 euphTyrcysAlaValaAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
2703 TTGCCCTTTGGCTTGTGTGCTGCTGCTGCGCAAAATCATGATTGTGG 2752
70 AlaSerAlaSerArgLysIleValProGlyAspAlaGlyValGlyIle 86
2753 GCGGCGGCTTGGCGCAACGCTGCGGCGCGCACCAAAACCATCA 2802
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyIle 103
2803 GCCGCTGGAAACGCTGCTGTTAAGCGGTACATGCTGCGCGACGGCGAC 2852
103 hrValAlaGlyGluIleuIleuIleuLeuAspProPheAlaGlyGly 119
2853 ATGTGAACAGGAGAAACGCTGCGGCACTGAGCTGTGGAAACAGAC 2902
120 ValAspVal 122
2903 AGCGATGTC 2911

seq_name: gb_ba:NM622491
seq_documentation_block:
LOCUS NM622491 326301 bp DNA linear BCT 04-DEC-2000
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 6/7.
ACCESSION AL162757 AL157959
VERSION AL162757.2 GI:7380371
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis Z2491.
Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 326301)
Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Felwell,T., Hamlin,N.,
Holroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skellern,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
Nature 404 (6777), 502-506 (2000)
JOURNAL
20222556
TITLE
2 (bases 1 to 326301)
Parkhill,J.
Direct Submission
Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT
Notes:
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source
1..326301
/organism="Neisseria meningitidis Z2491"
/db_xref="taxon:122587"
/contig="serogroup: A"
/contig="serogroup: A"
196..311
/contig="REP 2; hmms hit to HMM REP 2 (1 - 109), score:
87.43"
434..453
/contig="> 90% match to ATCCNNNNNNNGGAAT"
/label="GRS3
498..560
repeat_unit
repeat_unit
repeat_unit

```

seq_name:	gb_ba:AL646077				
seq_documentation_block:		189050 bp	DNA	linear	BCT 07-DEC-2001
LOCUS	AL646077				
DEFINITION	<i>Ralstonia solanacearum</i> Gm11000 megaplasmid, complete sequence;				
ACCESSION	segment 2/11.				
VERSION	AL646077 AL646053				
KEYWORDS	AL646077.1 GI:17430642				
SOURCE	<i>Ralstonia solanacearum</i> .				
ORGANISM	<i>Ralstonia solanacearum</i> . Bacteria; Proteobacteria; beta subdivision; <i>Ralstonia</i> group; <i>Ralstonia</i> .				
REFERENCE	1 (bases 1 to 189050)				
AUTHORS	Saranonbat,M., Genin,S., Attiguenave,F., Gouzy,J., Mangenot,S., Ariat,M., Billault,A., Brothier,P., Camus,J.C., Catalicio,L., Chandler,M., Choise,N., Claudel-Renard,C., Cunac,S., Demange,N., Gaspin,C., Lavie,M., Moisan,A., Robert,C., Sautin,W., Schlex,T., Siguler,P., Thebaul,P., Whalen,R., Wincker,P., Levy,M., Weissenbach,J. and Boucher,C.A. Genome sequence of the plant pathogen <i>Ralstonia solanacearum</i> unpublished 2 (bases 1 to 189050) Boucher,C.A.				



predicted by Framed"

/codon\_start=1  
/transl\_table=1  
/evidence=not\_experimental  
/product="HYPOTHETICAL PROTEIN"  
/protein\_id="CAD17330.1"  
/db\_xref="GI:17430647"  
/translation="MEANADPRMOPMRAVGRGATATLSALAPALADAPKPPMO  
DEPARAVSEAPQEVVSGTHPLVLIHADPAMVAQVSTDFVHNRVAPKSGRLI  
GKELRYNAREIWIQIIPASGTLRLDPLQATMRDGSAGVADLSPGALLGAMTS  
EPFVPH"

gene  
complement(8066..8488)  
/gene="RSP0180"  
/note="RSD04692"  
complement(8066..8488)  
/gene="RSP0180"  
/function="miscellaneous; hypothetical/global homology"  
/note="Product confidence : hypothetical  
Gene name confidence : hypothetical  
predicted by Codon\_usage  
predicted by Framed"  
/codon\_start=1  
/transl\_table=1  
/evidence=not\_experimental  
/product="CONSERVED HYPOTHETICAL PROTEIN"  
/protein\_id="CAD17331.1"  
/db\_xref="GI:17430648"  
/translation="MGAWSHASFGNDADIDFAEVAFTNDLSRVESAFDVAACDD  
LEAPEASQVAEAEEVARIQHWGVRNADSSIDAMIERITGITPRDVIHAKRALDR  
IVCEPSELLEIODESDEDKRDVARELRARLDGAPAR"  
complement(8552..8956)  
/gene="RSP0181"  
/note="RSD04693"

gene

alignment\_scores:  
Quality: 218.00 Length: 113  
Ratio: 2.868 Gaps: 1  
Percent Similarity: 67.257 Percent Identity: 41.593

## alignment\_block:

US-09-484-577A-4 x AL646077/rev ..

Align seg 1/1 to reverse of: AL646077 from: 1 to: 189050

12 ArgGluValValAlaValAlaGlyLysThrArgAspLeuLeu..... 26  
|||||  
163572 CGGGATGCTTCGCCGACACTAGTGAACGACGCGCGCTGAGCCTGCC 163523  
27 .....PhleuProAlaAlaLeuGluLeu 35  
163522 CCGCTTCGAGGTGACAGCGCGGATTCCTGCCGCGCGCTGCGCTGC 163473  
35 AlGluThrProProSerProThrAlaArgLeuThrAlaAlaLeuLeuAla 51  
163472 AGGCGCACCGCTCGCTCGCGCGCGGCGGATGGTGGCGCCGATCCTGATG 163423  
52 AlaLeuThrArgCysAlaValAlaAlaTrrAlaGlyLeuGlyArgLeuAsp 68  
163422 CTGCTGATCTGGTGGTCTGCTGCTGCTGCTGCTGCGCAAGTGCAT 163373  
68 eValAlaSerAlaSerArgLysLeuAlaProGlyAspArgValLysLeu 85  
163372 CATGCTCAACGCGGCGGCAAGATCATCCGTCGACGCGGCAAGACCA 163323  
85 AlGlnProLeuGluValGlyValAlaArgAlaThrAlaValArgAspGly 101  
163322 TTGCGCGCGGTGAGTGGCCACCGTGGCACAAGCTGATGTGAGAGAGGA 163273  
102 GlnThrValLysAlaGlyLysLeuLeuLeuLeuLeuLeu 114  
163272 CAGATCGTCAGAGCGGCGGATGTGCTGATCGACGCTGAC 163224  
seq\_name: gb\_ba:SS051197

## seq documentation block:

LOCUS SSU51197 28804 bp DNA linear BCT 16-MAY-1996  
DEFINITION Sphingomonas S88 sphingan polysaccharide synthesis (spsG), (spsS),  
(spsR), glycosyl transferase (spsQ), (spsI), glycosyl transferase  
(spsK), glycosyl transferase (spsL), (spsJ), (spsF), (spsD),  
(spsC), (spsE), Urf 32, Urf 26, ATP-binding cassette transporter  
(atrd), ATP-binding cassette transporter (atrb),  
glucosyl-1-isoprenylphosphate transferase (spsB), glucose-1-phosphate  
thymidyl transferase (rhaS), dTDP-6-deoxy-D-glucose -3,5-epimerase  
(rnsC) dTDP-D-glucose 4,6-dehydratase (rhaB),  
dTDP-6-deoxy-L-mannose-dehydrogenase (rnsD), Urf 31, and Urf 34  
genes, complete cds.  
ACCESSION U51197  
VERSION U51197.1 GI:1314561  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLES  
JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLES  
JOURNAL  
FEATURES  
source  
1. 28804  
/organism="Sphingomonas sp. S88"  
/strain="S88"  
/db\_xref="taxon:46624"  
complement(325..1944)  
/gene="spsG"  
complement(325..1944)  
/gene="spsG"  
/function="sphingan polysaccharide synthesis"  
/codon\_start=1  
/transl\_table=1  
/evidence=experimental  
/product="unknown"  
/protein\_id="AAC44055.1"  
/db\_xref="GI:1314562"  
/translation="MATSARVSGGLAVVOPIRARIVFPARAVLYTHPQSG  
LVTLGRFGLAVIVFVSQVGLSLILPSPMLGAAFLALALVIALPEADKAP  
TKLLRCYVLFMILVIMPNVLSVAGLPWISFRLEFVLTLLVCSISPFR  
DMAVTLASPMILAREMLGFLIIQIILISISQSPSIGRLVINSILVSPFATLMCL  
GSTRTPENMVNMLGCVSLVNIIGFEERAEHVLVGHIPSMQINDPSVILHTSQ  
VGTGYVNVNPGNPLVNGELVAMTPEFALHVNRSVSAALTLIFDLAMVLSAVS  
GRLSVNGVAVHATYLLMLGLRMTKKEGGLVGIISTMMYPAFAVYLVAMVAV  
HNRVLGGATGASNDGRHEDIRKGEPLKARPYFGTGPQSAEYVGRSQSGTSLSDS  
GELSTANDVGVFGLFAFYGTWVGAQLLRAGLTDSQGPLHLAVATVVVLTNRL  
VLSQTDNDVGVFGLFAFYGTWVGAQLLRAGLTDSQGPLHLAVATVVVLTNRL  
complement(1955..3313)  
/gene="spsS"  
complement(1955..3313)  
/function="sphingan polysaccharide synthesis"  
/codon\_start=1  
/transl\_table=1  
/evidence=experimental  
/product="unknown"  
/protein\_id="AAC44056.1"  
/db\_xref="GI:1314563"  
/translation="MWSGLPPEGPCPKAKOTGSPVGOAENVTRSKLHLGDKQWVG  
FGVKGGAVTSEFAFWLVARAFGPAVGAFGLATTAQMLVLTIGLDTILVIVAG  
CVSONRTGAVATVRAHATVAVAGGLVLMVMIILFQOQIATNLSGPPAAMPDLINGF  
LIPVSACRLASTTTRMGMDIGKQWMDPLGTFTLGAMLAGALLIGVAVHNPPLPSVL

```

gene
CDS
    complement(313..5325)
    /gene="spsr"
    /complement(313..5325)
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAC44057.1"
    /db_xref="GI:1314564"
    /translation="MPDILVRNOSSELDAAIRKAGGERTKIAAGYTSITVTTFPTT
    PYTISLASKRPVNTLTLOITINSSNLVKNVYNAQDFOPADQYNTANRTTGSNNVID
    CYTISGGTGVDTKSKMGILLRESKGVITKNSVDHFGALNNVLTVDGVTTLQGNFHD
    NRBDHTRNREASNVVINDNMFTNLYPVNGEHPDAIQFMTNGTKGNTITITNNYMG
    GDGLAFQGVFMGNESIPEYNTITKNNLYONGFTGHNVLGNKNTVDSNTIIKMD
    ALADWIRLDGVYGVKNTNNVDOVVITATSGITQSNNAVLATDAVTLRKIGDLMAGS
    KQALANLYSGVGYOPAGSAFAQVAKQAAQSTGANILLDLOFTATGIVDQTRM
    STDETAKLDLTAISNGMFOVCTGTGFEITRNSQIRPALPFTLNFDMKRTATGAV
    GOIMGVNOSWGISLADGEELVFTVNAAGQSTVATSGAKMTDTHKIKALTYDSAKG
    KAITIVDVVKSQGITVGSSTRAVEYAGLYIGSPENAVSGSVGEIEMRDALASQIIL
    ALNAAATASQPTTVTKPSLISLASAVTILGTPTVTALPAATTGTLTGGSASTA
    SVAQSLAQAIALMNAAGTAAVSATGSLSVANSLFATKKLVQDFYHV"
    5526..6473
    /gene="spsq"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="glycosyl transferase"
    /protein_id="AAC44058.1"
    /db_xref="GI:1314565"
    /translation="MEASPTPDVSLIVVAHSAPIFGICIRGIAAAGTAHEILLID
    NGGGTEVAHVAEFPFHVRIIPSEGNIGGAGNNRCAHARAPRILLVDPALPRGAI
    DLVAFKAHPDAAMAGRSYEPNGOLDHNPILPVRADPVSISSSPMRGGLPA
    DATAGPEVINGGFMVYDARVWRITDSEDFGFLISEIDLPQIRIRKGSVLYDPA
    VGVVHDGGHGISPTRYLFLITGGMHARHFGHGVAVTGMLMANAAKVVYVIGSL
    LGRISPRRAAMNLRDAMSTVFGQPRMHWGWRDVT"
    complement(6558..7460)
    /gene="spsl"
    /complement(6558..7460)
    /gene="spsr"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAC44059.1"
    /db_xref="GI:1314566"
    /translation="MKKLIYTVAAVAAALVAGSGSGEGRKIDKGVVAVSDGETTF
    ELNAEVOATPOSTDRKLAQMALQRIEIKITAKTAREDKDKTVEFLIQRRADE
    LIITTLTRDKTAGIAPQPTDAEIAQTPDAHDPRAQRKITSIEQIVPEPSAKFBE
    FAPLTKLDQAKLTAAGVOFRAPQDLTAALPPEAKIAAPAEFLILPQOGV
    TANYITGTTIOPLAGDQARELANTLTERFSKADQDLNRLKKARESVYQGYSP
    POLGGGAAAAPAAPAPAAQ"
    7588..8646
    /gene="spsk"
    /7588..8646
    /gene="spsk"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="glycosyl transferase"
    /protein_id="AAC44060.1"
    /db_xref="GI:1314567"
    /translation="MGDMAEATVTEAKKPKLKMCLASGGHLRIILLESVWKEH
    DYFFVEDTALGRSLAKSHVALVDHYALQAKLGHRLRLGCGMWRRLRSLSIRKH

```

```

gene
CDS
    complement(9552..10940)
    /gene="spsj"
    /complement(9552..10940)
    /gene="spsj"
    /function="secretion of sphingan polysaccharide"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="unknown"
    /protein_id="AAC44062.1"
    /db_xref="GI:1314569"
    /translation="WAGSALKFLMPGRDVPEDGYEPLAATVPPRDAHAGGRP
    DPFPRASALDRPDRREERREIRARFALATFTPPQPADRSSFAGRLVFLARLI
    SAIESORSHVVLVYGERGIGTSLILVLDVDESSYIYSATCGANANFADLFRVLE
    DVLFLHRSVSPNAGEAGSGLADRLTPGDSDELADICATGCPVLITIDEXR
    VYDPRFOOVAELIKNLSDRSARQOLYIAGVSSNIOELVGYPSIRRNITGIPRLE
    ETEVEMALIBETASGLFDPDRUTRIHLILAGSYFARLCHHAAALDQGRITVD
    LGHLKRALDQALILETGRMSPRAVIEMRKFGVGRYDPLVALGASRSTDMGCOAV
    VELLGSEITTAAYQOELDALSRQLESDADQDERRRFNDLSIPVYLMVGRIRI
    DSGLEALAAV"
    11569..12867
    /gene="spsf"
    /11569..12867
    /gene="spsf"
    /function="sphingan polysaccharide synthesis"
    /codon_start=1
    /transl_table=11
    /evidence="experimental"
    /product="unknown"
    /protein_id="AAC44063.1"
    /db_xref="GI:1314570"
    /translation="MKPEPGTEPVKDVQOAVEVRLCDRVSRSCVYLALATATAIQ
    PALAQRAFTRPGSGSEROSIKHATQOLEYNDNVNLDPRITTSARGVIVASPLDS
    IVPRAFGOLYLAGTVGYRFYRYRTNFRRENIISLGGDQRIASCYVHGEYGRHLT

```

alignment\_scores:

Quality	Ratio	Length
215.00	2.905	97
Percent Similarity: 76.289	Percent Identity: 45.361	

alignment\_block:

US-09-484-577A-4 x SSU51197/rev ..

Align seg 1/1 to reverse of: SSU51197 from: 1 to: 28804

```

27 pheleuProAlaAlaLeuGluIleValGluThrProProSerProThrAl 43
|||||
18808 TTCCTCCGCGCGCGCGCGCGAGATCAATCGAAGCTCCGCTGCCGCC 18759
43 aargleuThrAlaAlaLeuLeuAlaAlaLeuGlyTyrCysAlaValAlaT 60
|||||

```



18758 GCGCCTACCGCGCGGTGATGTTGGCGGCTGGCGATCCACCGCCT 18709  
60 rpAlaIleuGIYArgIleAspIleValAlaSerAlaSerArgIle 76  
18708 GGTGGCAATCGCGCGGTGAGTGTGGCGCGACGCGCGCATC 18659  
77 ValProGIYAspArgValLysLeuValGlnProLeuGluValAlaVala 93  
18658 GCGCCGATCGCGACGACAAAGATCGTCGATCCCGGAAGCGGTATCGT 18609  
93 lArGAlaIleHisValArgAspGlyGlnThrValLysAlaGlyGluIle 110  
18608 CCGCGCATTCGTGTGGCGCGACGAGGTTTCAGAAAGCGCAGGTGC 18559  
110 euIleGluLeuAspProPheAlaGlyIYValaIa 123  
18558 TGATCAGCTCGATCCACCATGTGGCAGCGACGCGCG 18518  
seq\_name: gp\_pat:AR068625  
documentation\_block:  
DEFINITION AR068625 28804 bp DNA linear PAT 29-SEP-1999  
ACCESSION AR068625  
VERSION AR068625.1 GI:6000832  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 28804)  
AUTHORS Pollock,T.J., Yamazaki,M., Thorne,L., Mikolajczak,M. and Armentrout,R.W.  
TITLE DNA segments and methods for increasing polysaccharide production  
JOURNAL Patent: US 5854034-A 1 29-DEC-1998;  
FEATURES  
location/Qualifiers  
source 1..28804  
BASE COUNT 4974 a 9804 c 9230 g 4796 t  
ORIGIN  
alignment\_scores:  
Quality: 215.00 Length: 97  
Ratio: 2.905 Gaps: 0  
Percent Similarity: 76.289 Percent Identity: 45.361  
alignment\_block:  
US-09-484-577A-4 x AR068625/rev ..  
Align seg 1/1 to reverse of: AR068625 from: 1 to: 28804  
27 PheLeuProAlaAlaLeuGluIleValGlnThrProProSerProThrAl 43  
18808 TTCTTCGCGCGCGCGTCGAGATCATCGAACGTCGCGTCCGCCACCGC 18759  
43 aArGLeuThrAlaAlaLeuAlaLeuAlaLeuPheTyCysAlaValaIar 60  
18758 GCGCCTACCGCGCGGTGATGTTGGCGGCTGGCGATCCACCGCCT 18709  
60 rpAlaIleuGIYArgIleAspIleValAlaSerAlaSerArgIle 76  
18708 GGTGGCAATCGCGCGGTGAGTGTGGCGCGACGCGCATC 18659  
77 ValProGIYAspArgValLysLeuValGlnProLeuGluValAlaVala 93  
18608 CCGCGCATTCGTGTGGCGCGACGAGGTTTCAGAAAGCGCAGGTGC 18559  
110 euIleGluLeuAspProPheAlaGlyIYValaIa 123

18558 TGATCAGCTCGATCCACCATGTGGCAGCGACGCGCG 18518  
seq\_name: gp\_ba:AF273216  
seq\_documentation\_block:  
LOCUS AF273216 3526 bp DNA linear BCT 17-JUL-2001  
DEFINITION Rhizobium leguminosarum bv. viciae bacteriocin (rzca) gene, partial  
cds: ABC transporter RzcB (rzcB) gene, complete cds; and  
transporter component RzcD (rzcD) gene, partial cds.  
ACCESSION AF273216  
VERSION AF273216.1 GI:10952793  
KEYWORDS  
SOURCE  
ORGANISM Rhizobium leguminosarum bv. viciae.  
Rhizobium leguminosarum bv. viciae  
Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;  
Rhizobiaceae; Rhizobium.  
REFERENCE 1 (bases 1 to 3526)  
AUTHORS Venter,A.P., Twelker,S., Oresnik,I.J. and Hynes,M.F.  
TITLE Analysis of the genetic region encoding a novel rhizobiocin from  
Rhizobium leguminosarum bv. viciae strain 306  
JOURNAL Canadian journal of microbiology. 47 (6), 495-502 (2001)  
MEDLINE 21360352  
PUBMED 11467725  
REFERENCE 2 (bases 1 to 3526)  
AUTHORS Venter,A.P., Twelker,S. and Hynes,M.F.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAY-2000) Biological Sciences, University of Calgary,  
2500 University Drive NW, Calgary, AB T2N 1N4, Canada  
location/Qualifiers  
source 1..3526  
/organism="Rhizobium leguminosarum bv. viciae"  
/strain="305"  
/db\_xref="taxon:387"  
/clone="pRle306c"  
/note="extrachromosomal  
biovlar: viciae"  
1..76  
/gene="rzca"  
<1..76  
/codon\_start=2  
/transl\_table=11  
/product="bacteriocin"  
/protein\_id="AAG25077.1"  
/db\_xref="GI:10952795"  
/translation="YETIDGSKHIVADSLIYRPTNA"  
144..2351  
/gene="rzcb"  
144..2351  
/note="required for bacteriocin secretion"  
/codon\_start=1  
/transl\_table=11  
/product="ABC transporter RzcB"  
/protein\_id="AAG25076.1"  
/db\_xref="GI:10952794"  
/translation="MMDIQTASAGVAEVEIRADSPPTPATIADSGVAICAVAGYR  
IASRPTISRELATAPARADRLRAKATVAKARVAREKISRLATPAAPALASR  
GTFAYPAGIAGAGRRRLINPDIPSANVADELALITSEPTILVORRFRPASQONF  
GRWELPAIMRYRRARFHVLIASLVQITVALVPLFQVAVKVALHNSYSTLIVLV  
GLAAVGLFDVQVLYRTVALSTTNRIDVELORRLRLINPLSYFETRAVGIVAR  
IRELEIRNMLTGGGLFSGLDLITLIEFVLSYSKLAIVASIPFYVAIGFLIR  
PLKERIDKFERGAFSOOLVETVVGIOTLRASAVEPVSSQMEERLAAYRSPKA  
YMLAKGONAIOYVNTITSAILLRGAQVVDIGELSGALVAFNMAGOVSPILRLS  
QIMODPOVOVSITARSIDLNPQERPREVANSIPPKALIAFKSVNPFYSPDGQDL  
KQINFSIRGEVIGIVGSPGSKSTLTKLVQRYLPNGQVVDGDLAQVDPAMLR  
NIGVAVQENMLFNRTIHNNICMVNPAMRAAIAQMARSGADQFTAKLPRQYDLIR  
EGANISGGRORLAIARALATNPILIIDEATSLADYSERIILGNNMEIYGRFVI  
ITAHRLATYRHNRIIGKMDGRIVEGTHETTLARNGLYAHLMLQDTPIES"  
2348..>3526  
/gene="rzcd"  
2348..>3526  
/gene="rzcd"  
/gene="rzcd"



```

/note="required for bacteriocin secretion"
/codon_start=1
/transl_table=11
/product="transporter component Rzcd"
/protein_id="AAG25078.1"
/db_xref="GI:10952796"
/translation="MNQIVQPPRLPPKRLQVNEFLPALEILELPPSPRIRALI
WFLCFTAGALIMSYIGFDIVATAGKIOPTRGVKIOSIEVGKTIAPVSNKAVE
TGDILVEDLPTEARREVNLTLSIALREVRERBALAEVNAWORGDLMSGTRRIE
VPEFADISPAEIRREBELLYTDLFOIASTLDSLAORNOCPAAVKRYTEMYTAQRAI
VATLADRVAMBSNIYDIDISAGSRGVDAVEITROKEENTLAEIGORAEKATVTAAS
BCKAKITPVADNAKQAAASREIDKEQDLVKAKRLESMTIKSPIINGIVOTSITT
VGQVVTAGELMRIVDASLEIEAYLPNRDIGVSPQPAVIVKAEAPPETRYGILNG
T"

BASE COUNT      672 a      1151 c      1025 g      678 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
```

gene  
DSDGNDKLSPSDSNLKDTLEKVKHNLVTTSKKEKVTIONMREADPAKFPVVKAT  
KDEKIEEIGONGERITSKOVDLAIKNGNEITODELSKYVDNELLKHSNVTNSLD  
KLISVSATFSSNDSRNVLVAFTMLDSSLDFARAA"  
3399. .5525  
/gene="1ktb"  
/allele="1ktb1.1"  
3399. .5525  
/gene="1ktb"  
/codon\_start=1  
/transl\_table=1  
/product="1ktb"  
/protein\_id="AAL12767.1"  
/db\_xref="GI:15987892"

CDS  
/translation="MEANHORNDGLVATLTMIAOVHNISLNPDEIKHEDLKGSLT  
TAMLLAAKSLAKAHKIKETSRHLVLPALVWQDNKHELVVDPDNNRYLTNL  
BODAPQILISODEFEACYOGQILVITRSVVGQLVKDPFWTIPVATYKRIETLTI  
VSILOIETALITPLEFQVYVMDKLVHRSSTLNTITVALATYIIEIVLSGLTYS  
HSTRIDVELGAKLERHLSPISYENRRVGDVVARRELDQIRNPLTGMALTSVD  
LLESPFFAVMYVSPKLTIVILGSLPCLTMSIFISPLRRLEKFAKSDNOAFL  
VESYTAIMIKAMAVAPQMTDMDKOLASVSSPRVLAITIGQOGLDOKTAVI  
NMLGAHLVIGSDLSIGQILAFNMLSGOVIATYITLAQMODFOQVIGSYRLGVIN  
SPTEYOQKLSLPEIKGDISEKNIRKRPDAPTLNNVNEIRQGEVIGYGRSGSG  
KSTLTKLQREYIPENGQVLDIGHDLADPWLRRQIGVVLQDVNLNRSIRENIAL  
SDGPMERVIYAAKLAGADHFISELREGYNTIVGOGAGLSGGORAIARALVN  
PKLIFEDATSAIDSESHIIMONOKICOGRTVILLIHRISTVKNADRIIYMERGEI  
VROCKHRELLNSGLSYLHQLQLN"  
5537. .>6952  
/gene="1ktb"  
/allele="1ktb1.1"  
5537. .>6952  
/gene="1ktb"  
/codon\_start=1  
/transl\_table=1  
/product="1ktb"  
/protein\_id="AAL12768.1"  
/db\_xref="GI:15987893"

BASE COUNT 2311 a 1219 c 1432 g 1990 t  
ORIGIN

Alignment\_scores:  
Quality: 182.00 Length: 103  
Ratio: 2.493 Gaps: 0  
Percent Similarity: 70.874 Percent Identity: 35.922

alignment\_block:  
US-09-484-577A-4 x AF314503 ..

Align seg 1/1 to: AF314503 from: 1 to: 6952

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36  
::: :::::::::: :::::::::::::::  
5636 AGAAAAAAGAGAGAAAGTAAATTCTTACCGGACATTGAAGAACTGATGCA 5685  
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaI 53  
||||| :::::::::: :::::::::::::::  
5686 AACCCCGGCTTTTAAACCAACGCTCTATTCTTATTGATTGATGCTAT 5735  
53 euPheTyrCysAlaValAlaTrrpAlaGlyLeuGlyArgIleAspIleVal 69  
::: :::::::::: :::::::::::::::  
5736 TTTTGGTTGTGCATTTGTGTCGACGTGAAGCAAGTGAATGCAATTGTCG 5785  
70 AlaSerAlaSerArgLysIleValProGlyAspArgValAlaLeuValG1 86  
||||| :::::::::: :::::::::::::::

5786 GCGACTGTCTCCCGCTAAATTACTTTAGTGGCAGAGTAAGAAATTAA 5835  
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyG1nt 103  
::: :::::::::: :::::::::::::::  
5836 ACCGATTGAAAACGCCATTGTACAGAAATTTTCGTTAAAGATGGGACGT 5885  
103 hrValysAlaGlyGluIleLeuIleGluLeuAspProPheAlaGlyGly 119  
::: :::::::::: :::::::::::::::  
5886 TTGTGGAAAAAGGCATTTATTGTCAGCTTAACGTGATGGGTTGAT 5935  
120 ValAspVal 122  
::: :::::::::: :::::::::::::::  
5936 GCGAGATATC 5944

seq\_name: gp\_ba:AF314506

seq\_documentation\_block:

LOCUS AF314506 6952 bp DNA linear BCT 21-DEC-2001  
DEFINITION Mannheimia haemolytica strain PH8 1ktcABD operon, partial sequence.  
ACCESSION AF314506  
VERSION AF314506.2 GI:15987894  
KEYWORDS

SOURCE

Mannheimia haemolytica.

ORGANISM

Mannheimia haemolytica; gamma subdivision; Pasteurellaceae;  
Bacteria; Proteobacteria;

REFERENCE

1 (bases 1 to 6952)  
Davies, R.L., Whittam, T.S. and Selander, R.K.

AUTHORS

Sequence diversity and molecular evolution of the leukotoxin (lktA)  
gene in bovine and ovine strains of Mannheimia (Pasteurella)

TITLE

gene in bovine and ovine strains of Mannheimia (Pasteurella)

JOURNAL

J. Bacteriol. 183 (4), 1394-1404 (2001)

MEDLINE

21101823

PUBMED

11157953

REFERENCE

2 (bases 1 to 6952)  
Davies, R.L., Campbell, S. and Whittam, T.S.

AUTHORS

Mosaic Structure and Molecular Evolution of the Leukotoxin Operon  
(lktCABD) in Mannheimia (Pasteurella) haemolytica, Mannheimia

TITLE

glucosida, and Pasteurella trehalosi

JOURNAL

J. Bacteriol. 184 (1), 266-277 (2002)

PUBMED

11741868

REFERENCE

3 (bases 1 to 6952)  
Davies, R.L., Whittam, T.S. and Selander, R.K.

AUTHORS

Direct Submission  
Submitted (18-OCT-2000) Institute of Biomedical and Life Sciences,  
Division of Infection and Immunity, University of Glasgow, Joseph

TITLE

Black Building, University Avenue, Glasgow G12 80Q, Scotland

JOURNAL

4 (bases 1 to 6952)  
Davies, R.L., Whittam, T.S. and Selander, R.K.

REFERENCE

Submitted (04-OCT-2001) Institute of Biomedical and Life Sciences,  
Division of Infection and Immunity, University of Glasgow, Joseph

AUTHORS

Black Building, University Avenue, Glasgow G12 80Q, Scotland

TITLE

Sequence update by submitter

JOURNAL

On Oct 9, 2001 this sequence version replaced gi:11762023.

COMMENT

Location/Qualifiers

FEATURES

1. .6952  
/organism="Mannheimia haemolytica"

gene

/strain="PH8"

gene

/serotype="A1"

gene

/db\_xref="taxon:75985"

CDS

/gene="1ktcABD operon"

gene

/gene="1ktc"

gene

/allele="1ktc1.1"

gene

/gene="1ktc"

gene

/codon\_start=1

gene

/transl\_table=1

gene

/product="1ktc"

gene

/protein\_id="AAL12769.1"

gene

/db\_xref="GI:15987895"

gene  
/translation="SLHKEMSCELLARVNIIPALENDQVMLIDNGIPAIACSWADLNL  
ETREVKYIKDINSLEPEWQSDGRWIIDWVAPGHSQLYKKMKOCYKPDIMVIRISIRFY  
PKQKELGTAIFYRGKLDKRAKRFDTYQELATALKNEFNTRK"  
/gene=".3324"  
/allele="1kta1.5"  
463..3324  
/gene="1kta"

CDS  
/codon\_start=1  
/transl\_table=11  
/product="leukotoxin"  
/protein\_id="AAG40290.1"  
/db\_xref="GI:11762024"

gene  
/translation="MGRLLTTLTSGNLKNTLTKTSGLHKAQSLTQAGSSLTGAKKI  
ILYIPONTYDTEQNGLDLVKAAEELGEVOREERNNIATAQTSIGTQAIIGTE  
RGIVASAPIDKLOKTAKGALGSAESIVONAKATYLSGISITGSLVGMDE  
ALONNSNOHALKAGIELNSLIENTANSVKTLDEGEDEISOGSKLONKRLGTGD  
KLNIGGDLKAGLGLDVISGLSGATPAALVADKANSTAKKAGFELANOVYGNITK  
AVSYLIAORVAGSSSTGPVAAALASTYSALISPLAFGIADKFNHAKSLSEVARE  
KUGITGDNLAEYORGTGTIDASVATINTALAAAGVSAASAASVIALPALLVSG  
ITGVISTLIQTSKQAMFEHVAANKINKITVEMKNNHKKVFPNGYDARYLANIDNMK  
FLNLNKELOARVIAITQOCWNNIGDLAGISRGKVLSGKAYVDAEELKHKIKAD  
KLVLDSNANGIIDVNSNGAKTOHILFRTPLTPGHEHREVOTGKVEYITKLNIRV  
DSMKITDGAASSTPLTNVORIGIELDNAGVATKTKETIKAKIGEGDNVFGSGT  
TEIDGEGVDYRHYRGNVCAITTDATKETEGGSTVNFNFEVTKALHEVISTHALV  
GNREKIEYRHSNNQHNAGYTKDTLKAVEETIGTSHNDFRSGKFNHAFNGDQVDT  
IDNGNDNRDLRFGKGDILLDGGNDLIDGGKNDLIDGGKDDIFVHKKGNGDIT  
DSDGNDKLSFSDSKDLTFEKKVKNLVTNSKKEKVTIDQFREADPAKEVPNKAT  
KDEKLEETIGONGERTSKQVVDLIAKNGKKTIDQELSKVVDNYELLHSHKNVTNSLD  
KLISVSATFSSNDNRNVLAFTSMLODLSLSLQPARAA"

CDS  
/gene="1kta"  
/allele="1kta2.1"  
3399..5525  
/gene="1kta"  
/allele="1kta"  
/codon\_start=1  
/transl\_table=11  
/product="1kta"

gene  
/protein\_id="AAL12770.1"  
/db\_xref="GI:15987896"  
/translation="MEANHQRNDLGLVALTMIAQYHNISLNEPEIKHFKFDLGGKLSL  
TAMLLAAKSLKAKHKEIKSRLHLVNLPAVMDNKGKFLVAVDNNRYLTYNL  
EODAPQLISODEFACQGOGLIIVTSRASVYGOALKEPFWETPAVIRKRIPELT  
VSILOJFALITPLEFOVYMDKLVYHRSFNLITVAVIYIPEIYLSGLRTYES  
HSTRIDVEIGAKLFRHLISPIYFERNRQDVARRELDQIKNEFTGQALTSVLD  
ILSFIFFAVMMITSPKLLVILSLPCYILMSLIFLISPLRRDEKPARSADNOAFL  
VESVATINMTKAMAVAPQMTDMDKOLASVSSSRVYVATIGQGOVLOKTYMTI  
NLWGAHLVYSGDLSIGLAFNMISGOVIAVIRIOLMODPOOVGIGSVRLQDVLN  
SPTEYOYKLSLPEIKGDISFKNIREFRKPDAPTLINNVLITROGEYITVGRSGC  
KSLTKLLOREYIPENGQVLDGHDLADPNMLPROGVYLDQNVILNRSIRENIAL  
SDPMPMERIYAKLAGAHDFISELRGIVTYGEGAGLSGGORQRIATARALVNN  
PKILIFPEATSALDYESSHIIIMQNMOKICQGRVILIAHRLSTVKNADRIIMKEGEL  
VEQGHHELLQNSNGLYSYLHQLQLN"

CDS  
/gene="1ktd"  
/allele="1ktd4.1"  
5537..>6952  
/gene="1ktd"  
/codon\_start=1  
/transl\_table=11  
/product="1ktd"

gene  
/translation="MKTWLSGTYEFLRYKNIAEYWKIRKELDHPNRKDESEFLPA  
HLELIETPVSKRPRLVAILMLFLVAIVLASVKEIVATAPKLTFSGRSKETKIP  
ENAIQVEIFVKDQFVKQGLVSLFALGSDADIKTIASLSLAKESYRQTLITAI  
EKESLPIVLDSTEFKDSSEEDRLRIKHLIEEYTTWOKOTKOTLAKYKREAEKORT  
FAYVRKEGATRIEOKLIDFKALYKOKSLKHELLAENKLIENONELAVYSKINE  
LENDLNVKKELELITQEFKSDYLEKIKOYIENROIRLEKKNOKQASMTIRAPIS  
GYOOLKITHYGVVTAETMLIIVPEDVLEATLVQNKDIGVAAQGEVITIKVET  
PYTRKGLITRIRIKHISPDATIEQPNVGLVFNATMTIDKKNLTPDGRKIDLSGWTITA  
EIKTGESVMSYLSPLSESVT"

BASE COUNT 2317 a 1211 c 1425 g 1999 t

## ORIGIN

alignment\_scores:  
Quality: 182.00 Length: 103  
Ratio: 2.493 Gaps: 0  
Percent Similarity: 70.874 Percent Identity: 35.922

alignment\_block:  
US-09-484-577A-4 x AF314506 ..

Align seg 1/1 to: AF314506 from: 1 to: 6952

```

20  LysThrArgAspGluLeuAlaPheLeuProAlaLeuGluIleValG1 36
   ::  :::::|||||  |||||  |||||  |||||  :::::
5636 AGAAAAAAGCGAAGATTTTACGGCACATTTAGAACATGATTGA 5685
   :::::  |||||  |||||  |||||  |||||  |||||
36  uThrProSerProThrAlaArgLeuThrAlaAlaLeuLeuAlaAla 53
   |||||  |||  |||||  |||  ||:::
5686 AACCCGGTTTCTAAAAAACACGCTCTAATTGCTTATTGATTTCCTAT 5735
   :::::  ||:::  ||:::  ||:::  ||:::  ||:::
53  eupheTyrcysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
   :::::  ||:::  ||:::  ||:::  ||:::  ||:::
5736 TTTTAGTGTGGCAATTGCTGCTCCAGTGAACCAAGTTGAATTGTG 5785
   :::::  ||:::  ||:::  ||:::  ||:::  ||:::
70  AlaserAlaserArgLysIleValProGlyAspArgValLysLeuValG1 86
   ||:::  |||||  |||||  |||||  |||||  |||||
5786 GCGACTCTCCCGGTAAATTACTTTTAGTGGCGAAGTAAGAAATTA 5835
   :::::  ||:::  ||:::  ||:::  ||:::  ||:::
86  nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyGlnT 103
   ||:::  |||||  |||||  |||||  |||||  |||||
5836 ACCGATTGAAGAACGCAATGTGTACAAAGAAATTTTCGTAAGATGGCGAGT 5885
   :::::  ||:::  ||:::  ||:::  ||:::  ||:::
103  hrValLysAlaGlyGluIleLeuIleGluLeuAspProPheAlaGlyGly 119
   ||:::  ||:::  ||:::  ||:::  ||:::  ||:::
5886 TTGTGGAAAAAGGCAATTATTAGTCAGCTTAACATGCATTGGGTTCTGAT 5935
   :::::  ||:::  ||:::  ||:::  ||:::  ||:::
120  ValAspVal 122
   :::::  ||:::  ||:::  ||:::  ||:::  ||:::
5936 GCAAGATATC 5944

```

Tue Jul 16 05:46:30 2002

us-09-484-577a-4.rge

---



gb_est2:BG831234	-	77.00	134.8
------------------	---	-------	-------

—

136 G 136



```

11:::||||| 11111  ::::| 11 11
246 TCGGTGCTGTCG...GGGCGC...TCCTGCTCGCTCTCTCTCTCTG 203
75 LysileValProGlyAspArgValLysleuValGlnProLeuGluValG1 91
202 CGGTGCTGCTCG...TGTGCTGCTCGCCAGGCTCTCGCCCATGG 159
91 yValValArgAlaThrHisValArgAspGlyGlnThrValLysAlaGly 108
158 CGGGGCTCGCTGCTCGCTCGCTCGCGCGCTCTCTGCTGCTAGTA 109
108 Luileu 110
108 CGCTGCTG 101

```

seq\_name: gb\_est2:BF781847

seq\_documentation\_block: 961 bp mRNA linear EST 12-JAN-2001

LOCUS BF781847 602107387F1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:4235732

DEFINITION 5', mRNA sequence.

VERSION BF781847 GI:12086883

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/.

TITLE 1 (bases 1 to 961)

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LINL at:

http://image.llnl.gov

Plate: L1AM9842 row: n column: 21

High quality sequence stop: 649.

Location/Qualifiers

1..961

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4235732"

/lab\_host="DH10B (TI phage-resistant)"

/note="Organ: Kidney; Vector: pCMV-SPORE6; Site\_1: NCI; Site\_2: Sall; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library. |"

BASE COUNT 249 a 257 c 285 g 170 t

ORIGIN

alignment\_scores: Quality: 87.00 Length: 98

Ratio: 1.611 Gaps: 4

Percent Similarity: 55.102 Percent Identity: 27.551

alignment\_block: US-09-484-577a-4 x BF781847 ..

Align seg 1/1 to: BF781847 from: 1 to: 961

29 ProAlaAlaLeuGluIleValGluThrProProSerProThrAlaArg.. 44

||||| ..:||||| ||| ||| ::|||

33 CCGGCAAGCGCTCAGCATCTGATCCAGACCTTCCAAAGGAGAGAA 82

```

45 .....LeuThrAlaAlaLeuLeuAlaAlaLeuP 54
83 GGTGGACACAGAGTCTCTGCTCTCTCCACACATTCATTTATGCTTC 132
54 heTyrcysAlaValAlaTrpAlaGlyLeu.....GlyArg 65
133 TCAGCTCATTAACAGGCTCTGCTCTCTGCCAGATCATCCAGGGGACG 182
66 IleAspIleValAlaSerAlaSerArgLysIleValProGlyAspArg 82
183 CTGACAGCCTGCTCGCCAGAGTGAAGATTCGTGGAAGCATF...GC 229
82 LysLeuValGlnProLeuGluValGlyValAlaArgAlaThrHisVal 99
230 TCAGCTGTGCTCGAGCGGAGTATATC.....CACATCT 261
99 rgaAspGlyGlnThrValLysAlaGlyGluIleLeuIleGluLeu 113
262 GCGATGCTCGCAGAGAGTACGCGGAGTGTGACCCACATG 305

```

seq\_name: gb\_gss:P284R

seq\_documentation\_block: 455 bp DNA linear GSS 16-MAR-2000

LOCUS P284R 455 bp DNA linear GSS 16-MAR-2000

DEFINITION Leishmania major Friedlin PAC P284 right end sequence, similar to

KPOXD J03885 K.pneumoniae oxalacetate decarboxylase. . . N=271,

Prob=1.0e-11; SW:ODP2\_ALCEU Q59098 D1HYDROLIPOAMIDE. . . N=122,

Prob=3.1e-16, genomic survey sequence.

VERSION AL160620 GI:7259004

KEYWORDS GSS.

SOURCE Leishmania major.

ORGANISM Leishmania major.

REFERENCE Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Leishmania

1 (bases 1 to 455)

AUTHORS Ivens,A.C., Lewis,S.M., Bagherzadeh,A., Zhang,L., Chan,H.M. and

Smith,D.F.

TITLE A physical map of the Leishmania major Friedlin genome

JOURNAL Genome Res. 8 (2), 135-145 (1998)

REFERENCE 98146435

2 (bases 1 to 158)

AUTHORS Taylor,R.G., Huckle,E.E.J., Ivens,A.C., Rajandream,M.A. and

Barrell,B.G.

TITLE Direct Submission

JOURNAL Submitted (14-MAR-2000) Leishmania major Friedlin genome sequencing

project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: barrelesanger.ac.uk and

allicatesanger.ac.uk

see http://www.edl.ac.uk/parasites/leish.html

details of Leishmania sequencing at the Sanger Centre are available

at http://www.sanger.ac.uk/Projects/L\_major/

The primer sequence can be obtained from allicatesanger.ac.uk.

Location/Qualifiers

1..455

/organism="Leishmania major"

/strain="Friedlin"

/db\_xref="taxon:5664"

/clone="PAC P284"

BASE COUNT 83 a 147 c 149 g 76 t

ORIGIN

alignment\_scores: Quality: 85.50 Length: 87

Ratio: 1.745 Gaps: 5

Percent Similarity: 56.322 Percent Identity: 33.333

alignment\_block: US-09-484-577a-4 x P284R ..

Align seg 1/1 to: P284R from: 1 to: 455



```

37  ThrProSerProThrAlaArgLeuThrAlaLeuAlaAlaLe 53
    ::::: ||||| ::::: ||||| ::::: |||
21  TCCCGGAGTCCCGCATGAGCAATGCAAGCGCCGTC.....GCCGG 64
    ::::: ||||| ::::: ||||| ::::: |||
53  UheTYrCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleValA 70
    ||||| ::::: ||||| ::::: |||
65  CCATGTCGCGCCGTC.....GACGTGCAG 90
    ::::: ||||| ::::: ||||| ::::: |||
70  lASerAlaSerArgIleValProGlyAspArgValIleValGln 86
    ::::: ||||| ::::: ||||| ::::: |||
91  TTGGCGGTACC.....ATGCGCGCGCGCATGCCGTATACCATCGAG 134
    ::::: ||||| ::::: ||||| ::::: |||
87  .....ProLeuGln.....ValGlyAlaValAla 94
    ::::: ||||| ::::: ||||| ::::: |||
135  TCGATGAGATGGAATCCCGTCGAGTCGAGCAGCGCGCGTCGTCA 184
    ::::: ||||| ::::: ||||| ::::: |||
94  gAlaThrHisValArgAspGlyInThrValIleGlyGluIleLeu 111
    ::::: ||||| ::::: ||||| ::::: |||
185  GGCCATCTCTGTCCACCCGCGCATGCGGTGCCGAGGCCAGGTGTGA 234
    ::::: ||||| ::::: ||||| ::::: |||
111  lEgIuLeuAsp 114
    ||||| ::::: ||||| ::::: |||
235  TCGAATCTCGAG 245
    ||||| ::::: ||||| ::::: |||

```

seq\_name: gb\_gss:AG163349

seq\_documentation\_block:

LOCUS AG163349 682 bp DNA linear GSS 09-JAN-2002  
DEFINITION Pan troglodytes DNA, clone: RP43-030A23.T7, genomic survey  
sequence.

ACCESSION AG163349.1 GI:16693027

VERSION GSS: GSS (genome survey sequence).

KEYWORDS Pan troglodytes male lymphocytes DNA, clone:lib:RP43-43 Chimpanzee

SOURCE

ORGANISM

Man troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Submitted (02-AUG-2001) Aseo Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimbes@gsr.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC  
end was generated during the R&D process and may have higher chance  
of clone tracking errors.  
PRIMERS  
Sequencing: T7  
LIBRARY  
Vector : pBAC3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI.  
Location/Qualifiers  
1. 682  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-030A23.T7"  
/sex="male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-43 Chimpanzee Male BAC library"  
183 a 138 c 138 g 221 t 2 others

alignment\_scores:  
Quality: 85.00 Length: 140  
Ratio: 1.164 Gaps: 6  
Percent Similarity: 52.143 Percent Identity: 30.714  
alignment\_block:  
US-09-484-577a-4 x AG163349 ..

Align seg 1/1 to: AG163349 from: 1 to: 682

```

3 SerLeuSerProArgMetLysSerAlaArgGluValAla..... 16
|||||
83 AGTCTCTACCCAGCGTGCATGCACTGCTGATCTTGCCTACTGCA 132
|||||
17 .....ValGlyGlyThrArgAspIleValAlaPheLeuProA 30
|||||
133 CCTCCACCTCTGCGTTCAATGATTCATGCTCAGCTCTCTAGTAG 182
|||||
30 lAlaLeuGluIleValAlaGluThrProSerProThrAlaArgLeuThr 46
|||||
183 CTGGGATT.....ACAGATGCTTGCACCATGCGCGCTAATT 220
|||||
47 AlAlaLeuLeuAlaAlaLeuPheTYrCysAlaValAlaTrpAlaGlyLe 63
|||||
221 TCTATATTATTAGTAGAGCGGGTTTCCCATGTTGTCAGCGCTGCT 270
|||||
63 uGlyArgIle.....AspIleValAlaSerAlaSerArgIleVal 77
|||||
271 TCAACTCTGACCTCATGTCATCCGCTGCTGCTGCTGCCAATGATTA 320
|||||
77 alProGlyAspArgVal.....LysLeuVal 85
|||||
321 CACTTATAGATAGACTCCAAAGTATATTAAAGATTGAGAAACTGTC 370
|||||
86 GlNProLeu...GluValGlyValAlaValAlaTrpHisValArgAspGl 101
|||||
371 CAGAACATCTTCAGATACATGATGATTAAAGACAGAAATGACTC 420
|||||
101 yGlnThrValIleValAlaGlyGluIle.....LeuIleGluLeuA 114
|||||
421 CCAAACTCCAGAAAGATGTGAGCTGTGCTACACCACTTATTACTTGT 470
|||||
114 sProPheAlaGlyGlyVal 120
|||||
471 GACCTTGGGAGGTAACCTTA 490
|||||

```

seq\_name: gb\_est:BG274115

seq\_documentation\_block:

LOCUS BG274115 432 bp mRNA linear EST 21-FEB-2001

DEFINITION WHE2231.G01\_N01Zs Aegilops speltoides anther cDNA library Aegilops

speltoides cDNA clone WHE2231.G01\_N01, mRNA sequence.

ACCESSION BG274115

VERSION BG274115.1 GI:13066099

KEYWORDS EST.

SOURCE Aegilops speltoides.

ORGANISM Aegilops speltoides

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

; Triticeae; Aegilops.

1 (bases 1 to 432)

Akhunov, E., Anderson, O.D., Chao, S., Chin, A., Choi, D.W., Close, T.J.,

Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Kianian, P., Lazo, G.R.,

Miller, R., Otto, C., Rausch, C.J., Seaton, C.L., Simons, K., Tong, J.C.,

The structure and function of the expressed portion of the wheat  
genomes - Anther cDNA library from Aegilops speltoides  
Unpublished (2000)  
Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 510595773

Fax: 5105595818  
Email: oanders@pw.usda.gov  
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
Seq primer: Stratagene SK primer.  
Location/Qualifiers  
1. 432  
/organism="Aegilops speltoides"  
/cultivar="F2 from 2-12-4-8-1-1-1 (1) x PI36909-12-811-(1)"  
/db\_xref="taxon:4573"  
/clone\_lib="Aegilops speltoides anther cDNA library"  
/tissue\_type="anther"  
/dev\_stage="Premeiotic anthers"  
/lab\_host="E. coli SOLR"  
/note="Vector: Lambda uni-ZAP XR, excised phagemid; Site-1: EcoRI; Site-2: XhoI; Plants were grown in a growth chamber at the University of California, Davis (Akhunov). Premeiotic anthers were harvested, total RNA and poly(A) RNA were prepared, from each tissue and then pooled, a cDNA library was made, and the cDNA clones were in vivo excised to give plasmid phagemids in the TU Close Lab (Akhunov, Chin, Choi, Close, Fenton, Khaman, Otto, Simons, Zhang) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson Lab (all other authors)."

BASE COUNT 81 a 128 c 134 g 89 t  
ORIGIN

alignment\_scores:  
Quality: 84.00 Length: 67  
Ratio: 1.867 Gaps: 3  
Percent Similarity: 67.164 Percent Identity: 38.806

alignment\_block:  
US-09-484-577A-4 x BG274115 ..

Align seg 1/1 to: BG274115 from: 1 to: 432

```

3 SerLeuSerProArgMetLysSerAlaArgIuValAlaAlaValGlyI 19
|||||
33 AGCTTGTCCCTTATGGCGCGCAAGACGCGTTGACCAAGCTCG 82
|||||
19 YLVTThArgAspLeuAlaPheLeuProAlaAlaLeuGluLeValG 36
|||||
83 AAGAACACGCGCTGAGCGACATTCCTCAGCGCGCGCTG.....GCCA 126
|||||
36 IuThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAla 52
|||||
127 GAGCGCCGCGACAGCCAC...AGCGTGAATCCATCAATATTTGGCA 173
|||||
53 Leu.....PheTYCYsAlaValAlaTTPAlaGlyLeuGlyArgI 66
|||||
174 ATGCCCATGCAATCATGTTTCATGCGATGGGGGCGCTGGCGGCT 223
|||||
66 e 66
224 G 224

```

seq\_name: gb\_gss:AQ991716

seq\_documentation\_block:  
LOCUS AQ991716 756 bp DNA linear GSS 14-AUG-2000  
DEFINITION Rf00179F Photorhabdus luminescens strain W14 M13 library  
ACCESSION AQ991716  
VERSION AQ991716.1 GI:9650212  
KEYWORDS GSS.  
SOURCE Photorhabdus luminescens.  
ORGANISM Photorhabdus luminescens  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE  
1 (bases 1 to 756)  
french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,  
Daborn,P.T., Bowen,D. and Blattner,F.R.  
A genomic sample sequence of the entomopathogenic bacterium  
Photorhabdus luminescens W14: potential implications for virulence  
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
20378633  
JOURNAL MEDLINE  
COMMENT Contact: french-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bsrfc@bath.ac.uk  
This is one of a selected subset of flipped clones from the M13  
library. For annotation of identified clones (BLASTX, BLASTN  
mapping to E. coli K12 genome) please see french-Constant et al.  
2000, Nucleic Acids Res.  
Seq primer: M13 Reverse  
Class: Shotgun.

FEATURES  
source Location/Qualifiers  
1. 756  
/organism="Photorhabdus luminescens"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone\_lib="PLG00179F"  
/clone\_lib="Photorhabdus luminescens strain W14 M13  
library"  
/dev\_stage="primary phase variant"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 185 a 205 c 132 g 214 t 20 others  
ORIGIN

alignment\_scores:  
Quality: 84.00 Length: 55  
Ratio: 2.270 Gaps: 2  
Percent Similarity: 67.273 Percent Identity: 38.182

alignment\_block:  
US-09-484-577A-4 x AQ991716/rev ..

Align seg 1/1 to reverse of: AQ991716 from: 1 to: 756

```

60 TTPAlaGlyLeuGlyArgIleAspIle...ValAlaSerAlaSerArgI 75
|||||
509 TGGCGACGATTAGCGCGCTGATTAAGAGGCGTGGCTTCGCGGTGCT 460
|||||
75 sIleValProGlyAspArgValIleValGlnProLeuGluValGly 92
|||||
459 GGTTCGATGTATTCGT...AAACAGTTCACGCCCGCCGCAACGGGA 413
|||||
92 alValAlaGAlaThrHisValaArgAspGlyGlnThrValIlyAlaGlyI 108
|||||
412 TTATCAGCCAGATTAAGTTGTTGAGGAGAAACATTAACCGCGGTGAG 363
|||||
109 IleLeuIleGluLeu 113
|||||
362 ACACCTGCTCCAGCTC 348

```

seq\_name: gb\_est2:BG758850

seq\_documentation\_block:  
LOCUS BG758850 973 bp mRNA linear EST 15-MAY-2001  
DEFINITION 602713239F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4653577 5',  
ACCESSION BG758850  
VERSION BG758850.1 GI:14069503  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 973)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LNCM1699 row: n column: 10  
 High quality sequence stop: 2.

## FEATURES

1. 973  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4853577"  
 /clone\_lib="NIH-MGC 48"  
 /tissue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: B-cells; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 232 a 261 c 279 g 199 t 2 others  
 ORIGIN

## alignment\_scores:

Quality: 83.50 Length: 77  
 Ratio: 1.815 Gaps: 3  
 Percent Similarity: 59.740 Percent Identity: 35.065

## alignment\_block:

US-09-484-577A-4 x BG758850 ..

Align seg 1/1 to: BG758850 from: 1 to: 973

```

49 LeuleuAlaAlaLeuphetrCysAlaValAlaTrpAlaGlyLeuGlyA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 CTGCTAGGAGGAGAACTTTACTCTTAGAGAGGAGGTGTCGACTCTGGCAG 522
65 gtleaspileval.....AlaserAlaserArgLysilevalProGlyA 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
523 GTTGTCAACTGTTATCCACGGCAGGAGGTAGTGGCTGGTGGGTGGT 572
80 spArg_VaLlysLeuValGlnProLeuGluValGlyValAlaArgAlaTh 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
573 CCGGTGCGACTTTTGTGTCGCCCATTTGGCTGGCTCTTACGCGTTCG 622
96 rhtsValArGAspGlyGlnThrValLysAlaGlyLuleuLeuIleGluL 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
623 GAGAGTCCGGGACGGGGGCGATCCCCCGT..... 650
113 euAspProPheAlaGlyLysValAspVal 122
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
651 .....GAGTATTCGCGCGAGGTGAGTGC 674
seq_name: gb_estc:BJ189099

```

seq\_documentation\_block: 456 bp mRNA linear EST 24-JAN-2002  
 LOCUS BJ189099  
 DEFINITION BJ189099 normalized full length cDNA library, chloronemata,

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 COMMENT

caulonemata and malformed buds *Physcomitrella patens* subsp. *patens*  
 cDNA clone pph45c13 5', mRNA sequence.  
 BJ189099.1 GI:18357040  
 EST.  
*Physcomitrella patens* subsp. *patens*.  
*Physcomitrella patens* subsp. *patens*.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
 Bryopsida; Funariidae; Funariales; Funariaceae; *Physcomitrella*.  
 1 (bases 1 to 456)  
 Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,  
 Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe  
 M.  
 Comparison of the moss *Physcomitrella patens* genome with flowering  
 plants genome  
 Unpublished (2002)  
 Contact: Tadasi Shin-I  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp  
 A backbone of the vector is pBluescript II, that was in vivo  
 excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI  
 digested-5' end of cDNA is ligated to SalI site of the vector, and  
 the BamHI digested-3' end including poly-A tail is ligated to BamHI  
 site of the vector. cDNA insert could be amplified with  
 conventional T7 and T3 primers. This normalized full-length cDNA  
 library was generated basically according to the method described  
 in Genome Research 10, 1617-1630 (2000). Carninci, P. et al.  
 Protonemata were blended by the POLYTRON, and then cultivated on  
 the BCD medium containing 0.5mM BA (benzylaminopurine) for 8 to 13  
 days under the continuous light.  
 Location/Qualifiers  
 1. 456  
 /organism="Physcomitrella patens subsp. patens"  
 /db\_xref="taxon:145481"  
 /clone="pph45c13"  
 /clone\_lib="normalized full length cDNA library,  
 chloronemata, caulonemata and malformed buds"  
 /tissue\_type="mixture of chloronemata, caulonemata and  
 malformed buds"

BASE COUNT 97 a 97 c 129 g 133 t  
 ORIGIN

## alignment\_scores:

Quality: 82.00 Length: 93  
 Ratio: 1.608 Gaps: 3  
 Percent Similarity: 54.839 Percent Identity: 26.882

## alignment\_block:

US-09-484-577A-4 x BJ189099 ..

Align seg 1/1 to: BJ189099 from: 1 to: 456

```

38 ProProserProthrlaArgLeu..... 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25 CCCCCCCTCCGACCTCCGCGCTGACTGTCATTCGTCGAGAGAG 74
46 .ThrlaAlaLeuleuAlaLeuphetrCysAlaValAlaTrpAlaG 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
75 TACAGTTTGTACATCCCTGACGATTTGTCTGCGACGTT.....C 115
62 lyeuGlyArgGtleaspilevalAlaserAlaserArgLysilevalPro 78
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 ATTATGTCGAGATTGACTTGCTGAGCGGCTGAGACCTTGTTAGAC 165
79 GlyAspArgValLysLeuValGlnProLeuGluValGlyVal..... 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 GTTGAAGATACACTTTCGCCACCCGTCGAGAACTTGGCAGCGGGTGTGA 215

```

84 LeuValGlnProLeuGluValGlyValValArg 94

264  
||||| :||  
mmcmctccacnccgggtgatgcttgcagaagcttg 262

224 .....TTCACCTGTGCTGAAGGGGGTGATGGTTGGAGAAGGTTG 262

67 .....AspIleValAlaSerAla 73  
 263 GTGCTTTTCATGCGACAGCATGCTCCGACATGACGATCTGTGG 312  
 73 rArgLysIleValProGly.....AspArgValLysLeuValG 86  
 313 GGGCTCTGTGTGCCCCCGGATTTTCAGAGGATCTCTTCAGCTTTC 362  
 86 InProLeuGluValAlaValAlaValAlaThrHisValAlaGspolylGln 102  
 363 AGACGTCTCGATGTGTGTATCATCAGCCCTCAGACCAAGCTCAGAGGAG 412  
 103 ThrValLysAlaGlyGluIleLeuIleGluLeuAspProPheAlaGly 119  
 413 TGCAGCTTCGAGGCGCCAGCTTAAAGCCTGTGATTTACAGGCTGTC 462  
 119 yValAsp 121  
 463 TGTGAC 469  
 seq\_name: gb\_est1:AL514080

seq\_documentation\_block:  
 LOCUS AL514080 907 bp mRNA linear EST 13-FEB-2001  
 DEFINITION AL514080 LFI\_NFL006\_PL2 Homo sapiens cDNA clone CIOBA004ZE11 5  
 ACCESSION AL514080  
 VERSION AL514080.1 GI:12777574  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 907)  
 AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.  
 FEATURES  
 source  
 1..907  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CIOBA004ZE11"  
 /clone\_lib="LFI\_NFL006\_PL2"  
 /tissue\_type="placenta"  
 /note="Vector: PCWSPORT 6; site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCWSPORT 6 vector. Library was normalized. Library was constructed by a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: liliang@life-tech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 192 a 236 c 273 g 194 t 12 others  
 ORIGIN

alignment\_scores:  
 Quality: 80.00 Length: 69  
 Ratio: 1.860 Gaps: 1  
 Percent Similarity: 62.319 Percent Identity: 36.232

alignment\_block:

us-09-484-577a-4 x AL514080/rev ..

Align seg 1/1 to reverse of: AL514080 from: 1 to: 907

40 SerProThrAlaArgLeuThrAlaLeuLeuAlaAlaLeuPheTyr 56  
 280 TCGCGGAGACCTCTGCTGCTGCGACCTGCTACTTCTTCTTCCCTG 231  
 56 sAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleValAlaSerAla 73  
 230 TGTCTCGGATTTGTTACTCTCGAGATTGCCCTGCTGCTGGAAGTA 181  
 73 eArgLysIleValProGlyAspArgValLysLeuVal.....GlnPro 87  
 180 GTGTAAGACACCCCTGCTGACACGACCGCAGCTCGGAGCACACCG 131  
 88 LeuGluValGlyValAlaValAlaThrHisValAlaGspolylGln 104  
 130 CTGCTCTCGCAATAGTCGATTCCTCCACACGCTCCGTGAGCCGAGT 81  
 104 lLysAla 106  
 80 CCGGCA 74

seq\_name: gb\_est2:BE481913  
 seq\_documentation\_block:  
 LOCUS BE481913 360 bp mRNA linear EST 28-AUG-2000  
 DEFINITION 167524 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION BE481913  
 VERSION BE481913.1 GI:9601446  
 KEYWORDS EST.  
 SOURCE cow.  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 360)  
 AUTHORS Sonstegard, T.S., Capuco, A.V., Van Tassel, C.P., Ashwell, M.S. and Wells, K.D.  
 TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Sonstegard TS  
 USDA, ARS, Beltsville Agricultural Research Center  
 Bldg. 200, Rm 2A, Beltsville, MD 20705, USA  
 Tel: 301 504 8416  
 Fax: 301 504 8414  
 Email: tads@psi.barc.usda.gov  
 Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross-match with the -minscore 18 PCR Primers  
 FORWARD: AGGAACAGCTATGACAT  
 BACKWARD: GTTTCACAGCAGCAGC  
 Plate: 19 row: 0 column: 4  
 Seq primer: ATTGAGTGACATATAC.  
 FEATURES  
 source  
 1..360  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="BARC 5BOV"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV SPORT6; site\_1: XbaI; site\_2: XhoI; library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states.."

BASE COUNT 73 a 123 c 107 g 57 t  
 ORIGIN

alignment\_scores:  
 Quality: 79.50 Length: 84  
 Ratio: 1.691 Gaps: 3  
 Percent Similarity: 55.952 Percent Identity: 32.143

alignment\_block:  
US-09-484-577a-4 x BE481913/rev ..

Align seg 1/1 to reverse of: BE481913 from: 1 to: 360

```

39 ProserProthrAlaArgLeuThrAlaLeuLeuAlaLeuAlaLeuPheTy 55
|||||.....:|||||.....:|||||.....:
319 CCGAGTCGGGTGTTCAAGTTTCGCCGAAATCTGTGCCGTCTCTGT 270
55 rcysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleValAlaSerA 72
:::.....:|||||.....:|||||.....:
269 AGATTCTGTGTGTGGATCCCAATACCTCCGCCCTCTCTCCATCCAC 220
72 laserArgLysIleValProGlyAspArgValLysLeuValGlnProLeu 88
:::.....:|||||.....:|||||.....:
219 GGCACCCCGCGGTGCTGCCCTTGATTCGGGGCGTCTGCGAACCCGTGT 170
89 GluValGlyValAlaArgAlaThrHisValArgAspGlyGlnThrValLy 105
|||||.....:|||||.....:|||||.....:
169 GAACCTGCGTGTCTC.....CACGTAGCCGACGCGCAT.....GA 135
105 sAlaGlyGluIleLeuIleGluLeuAspProPheAlaGlyValAspV 122
:::.....:|||||.....:|||||.....:
134 AGCGGGG.....CTCCCCAGAGCGGCGCGGACACGCGGTGTAGAAA 91
122 al 122
::
90 TA 89

```

seq\_name: gb\_est2:D27997

seq\_documentation\_block:

LOCUS D27997 394 bp mRNA linear EST 20-NOV-1995  
DEFINITION CELR007EIF Yui Kohara unpublished cDNA Caenorhabditis elegans cDNA  
clone yk7el 5', mRNA sequence.

ACCESSION D27997  
VERSION D27997.1 GI:522753

KEYWORDS EST  
SOURCE

ORGANISM

Caenorhabditis elegans.  
Caenorhabditis elegans.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

REFERENCE 1 (bases 1 to 394)  
Kohara, Y., Mitsuki, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and

AUTHORS

Toward an expression map of the C. elegans genome  
Unpublished (1994)

TITLE

JOURNAL

Contact: Yui Kohara

Genome Biology Lab.

National Institute of Genetics

Yata 111, Mishima, Shizuoka 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

Location/Qualifiers

1..394

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(e1489)"

/db\_xref="taxon:6239"

/clone="yk7el"

/clone\_lib="Yui Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue\_type="whole animal"

/dev\_stage="varied"

BASE COUNT 112 a 93 c 85 g 102 t 2 others

ORIGIN

alignment\_scores:  
Quality: 79.50 Length: 56  
Ratio: 2.650 Gaps: 1  
Percent Similarity: 53.571 Percent Identity: 33.929

alignment\_block:  
US-09-484-577a-4 x D27997 ..

Align seg 1/1 to: D27997 from: 1 to: 394

```

75 LysIleValProGlyAspArgValLysLeuValGlnProLeuGluVal.. 90
|||||.....:|||||.....:|||||.....:
86 AAATCAAGAGGAGCGACAGAGTTACCAAGACCAACCCCTATTGTGTCT 135
91 .....GlyValV 93
136 TTCCGCCATGAAATGGAGATGTCATGACTCTCCATTGCTGGACCG 185
93 alArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle 109
|||||.....:|||||.....:|||||.....:
186 TCAAGCTATCATGACACACACAGAACCAATGTAGTCCGCGAGACTTG 235
110 LeuIleGluLeuAspPro 115
:::.....:|||||.....:|||||.....:
236 GTAGTGAAGTTGACCA 253

```

Tue Jul 16 05:46:34 2002

us-09-484-577a-4.rst

Page 10





Percent Similarity: 71.845 Percent Identity: 49.515

Alignment\_block:

US-09-484-577A-4 x AA254335

Align seg 1/1 to: AA254335 from: 1 to: 1428

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
94 CGCAGCGCGGAGAACAGCGCTTTGGCCCGCATTTGGCACTGACCGCA 143
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
36 uThProProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
144 TACGCGGCTCTCTGCGCTCGCAATGGCGCGCGCTTTATATATGCGGT 193
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
53 eupheryCysAlaValAlaTrpAlaGlyLeuGlyArgGlyIleAspIleVal 69
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
194 TTGCGCTTTTGGCTTTGTGTGTGTCCTGCGTTCGCGCAAAATCGATATGTG 243
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
70 AlaSerAlaSerArgGlyIleValProGlyAspArgValIysLeuValG1 86
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
244 GCGCGGCTTTCGGCGGCAAAACGCTCGCGCGCGCGCGCAAAACCATCCA 293
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
86 nProLeuGluValGlyValValArgAlaThrHisValArgAspGlyGlnT 103
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
294 GCCGCTGGAACGCGCGGTGTTAAGCGGTACATGTGCGCGAGCGGCGACG 343
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
103 hrValIysAlaGlyIuIleLeuIleGluLeuAspProPheAlaGlyGly 119
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
344 ATGTGAACAGCGAGAAACGCTGCGGACGACTGAGCGCTGTGGGAACAGAC 393
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
120 ValAspVal 122
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
394 AGCGATGTG 402
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

seq\_name: /SID1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA81532

seq\_documentation\_block:

ID AA81532 standard; DNA; 24158 BP.

AA81532;

04-DEC-2000 (first entry)

N. meningitidis partial DNA sequence gnm\_79 SEQ ID NO:79.

Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
antigen; vaccine; diagnosis; infection; antibacterial; identification;  
Meningococcus B; Memb; ds.

Neisseria meningitidis.

MO200022430-A2.

20-APR-2000.

08-OCT-1999; 99WC-US23573.

09-OCT-1998; 98US-0103794.

30-APR-1999; 99US-0132068.

(CHIR) CHIRON CORP.

Frazier CM, Hickey E, Peterson J, Tettelin H, Venter JC;  
Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
Rappuoli R, Pizza M;

WPI: 2000-318079/27.

Isolated nucleotide sequences of Neisseria meningitidis which can be  
used in the diagnosis and treatment of N. meningitidis infection and  
other Neisserial infections, for example, N. gonorrhoea -

PS Claim 7; Page 1464-1471; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic  
CC proteins from Neisseria genomic sequences. AA81453 to AA82414  
CC represent specifically claimed Neisseria meningitidis genomic DNA  
CC sequences; AA81260 to AA81303 and AA825620 to AA825663 represent  
CC Neisseria DNA sequences and their corresponding proteins; AA81254 to  
CC AA81259 and AA81304 to AA81321 represent PCR primers used in the  
CC isolation of Neisseria meningitidis DNA sequences; and AA81322 to  
CC AA81452 represent Neisseria meningitidis MenB polynucleotide ORF  
CC sequences, which are all used in the exemplification of the present  
CC invention. The nucleic acid sequences, protein sequences, and antibodies  
CC against them, can be used in the manufacture of a composition. The  
CC composition can be used as a medicament (or in the manufacture of a  
CC medicament) for treating, preventing or diagnosing infection due to  
CC Neisserial bacteria. For example, some of the identified proteins could  
CC be components of vaccines against Meningococcus B; against all serotypes;  
CC and/or against all pathogenic Neisseriae. Identification of biological probes,  
CC particularly organism-specific probes. Attempts to make effcacious  
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.  
CC Multivalent vaccines have also been tried but none have successfully  
CC overcome antigenic variability. The provision of further, complete  
CC sequences may provide an opportunity to identify secreted or surface  
CC exposed proteins that may be presumed targets for the immune system and  
CC which are not antigenically variable or at least more conserved than  
CC other more variable regions.

Sequence 24158 BP; 5559 A; 6297 C; 6241 G; 6061 T; 0 other;

alignment\_scores:

Quality: 238.00 Length: 103

Percent Similarity: 3.216 Gaps: 0

Percent Identity: 71.845 Percent Identity: 49.515

alignment\_block:

US-09-484-577A-4 x AA81532

Align seg 1/1 to: AA81532 from: 1 to: 24158

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
17976 CGCAGCGCGGAGAACAGCGCTTTTGGCCCGCATTTGGCACTGACCGCA 18025
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
36 uThProProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18026 TACGCGGCTCTCTGCGCTCGCAATGGCGCGCGCTTTATATATGCGGT 18075
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
53 eupheryCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18076 TTGCGCTTTTGGCTTTGTGTGTGTCCTGCGCAAAATCGATATGTG 18125
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
70 AlaSerAlaSerArgGlyIleValProGlyAspArgValIysLeuValG1 86
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18126 GCGCGGCTTTCGGCGGCAAAACGCTGCGCGCGCGCGCAAAACCATCCA 18175
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
86 nProLeuGluValGlyValValArgAlaThrHisValArgAspGlyGlnT 103
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18176 GCCGCTGGAACGCGCGGTGTTAAGCGGTACATGTGCGCGAGCGGCGACG 18225
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
103 hrValIysAlaGlyIuIleLeuIleGluLeuAspProPheAlaGlyGly 119
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18226 ATGTGAACAGCGAGAAACGCTGCGGACGACTGAGCGCTGTGGGAACAGAC 18275
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
120 ValAspVal 122
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
18276 AGCGATGTG 18284
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

seq\_name: /SID1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAF21611

seq\_documentation\_block:

ID AAF21611 standard; DNA; 349980 BP.



CC sequences which overlap each other at the beginning and end of each  
 CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at  
 CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at  
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the  
 CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to  
 CC AAF21606 represent PCR primers which are used in the exemplification of  
 CC the present invention. The NMB genome and fragments from it have  
 CC antibacterial activity, and can be used in vaccines and gene therapy.  
 CC Neisseria nucleic acids, proteins and/or antibodies which binds to the  
 CC proteins can be used in compositions for treating or preventing infection  
 CC due to Neisserial bacteria or as a diagnostic reagent for detecting the  
 CC presence of Neisserial bacteria or of antibodies raised to Neisserial  
 CC bacteria. Computers, computer memory, computer storage medium or computer  
 CC databases can be used in a search to identify open reading frames (ORFs)  
 CC or coding sequences within the NMB genome. The DNA sequences provide  
 CC further opportunities to find antigenic or immunogenic proteins which are  
 CC more effective in vaccines than the outer membrane proteins currently  
 CC used.

Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

# alignment\_scores:

Quality: 238.00 Length: 103  
 Ratio: 3.216 Gaps: 0  
 Percent Similarity: 71.845 Percent Identity: 49.515

## alignment\_block:

US-09-484-577A-4 x AAF21612/rev ..

Align seg 1/1 to reverse of: AAF21612 from: 1 to: 349980

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluTleValG1 36
   ::::: ||| ||||| ||||| ||||| ||||| ||||| |||||
22829 CGCAGCGCGGAGAACAGCGGTTTGGCCCGGCAATTGGAACTGACCGA 23780
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaLal 53
   ::||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
23779 TACGCGCGCTCTGCTCCGAAATGGCGCGCGCTTTATTATAGGCGT 23730
   ::||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
53 euPheTyrCysAlaValAlaTrpAlaGlyLeuGlyArgTleAspTleVal 69
   ::||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
23729 TTGCGCGTTTGGCTTTGGTTGGTCTGTCGGCAAAATGAAATGATGTG 23680
   ::||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
70 AlaSerAlaSerArgTleValProGlyAspArgValLysLeuValG1 86
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23679 GCGCGCGCTTGGCGCAAAAGGTGCGCGCGCGGCGCAAAACATGCA 23630
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyLnt 103
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23629 GCCGCTGAAACGGCGGTGTTAAGCGGTACATGTCGCGACGCGGACG 23580
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
103 hrValLysAlaGlyLntLeuLeuLeuLeuLeuAspProPheAlaGly 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23579 ATGTGAACAGCAGGAGAAACGTCGCGCACTGAGGCTGTGGAAACAG 23530
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 ValAspVal 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23529 AGCGATGTG 23521

```

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAA81489

seq\_documentation\_block:

ID AAA81489 standard: DNA; 837096 BP.

AC AAA81489;

XX 04-DEC-2000 (first entry)

XX N. meningitidis partial DNA sequence gmm\_37 SEQ ID NO:37.

DE Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;

XX antigen; vaccine; diagnosis; infection; antibacterial; identification;

KM Meningococcus B; MenB; ds.

XX Neisseria meningitidis.

OS WO200022430-A2.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23573.

XX 09-OCT-1998; 98US-0103794.

XX 30-APR-1999; 99US-0132068.

XX (CHIR) CHIRON CORP.

XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,

XX Masiarani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

XX Rappuoli R, Pizzi M;

XX WPI; 2000-318079/27.

XX Isolated nucleotide sequences of Neisseria meningitidis which can be

XX used in the diagnosis and treatment of N. meningitidis infection and

XX other Neisserial infections, for example, N.gonorrhoea -

XX Claim 7; Page 629-865; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic

XX proteins from Neisseria genomic sequences. AAA81453 to AAA82414

XX represent specifically claimed Neisseria meningitidis genomic DNA

XX sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent

XX Neisseria DNA sequences and their corresponding proteins; AAA81254 to

XX AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the

XX isolation of Neisseria meningitidis DNA sequences; and AAA81322 to

XX AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF

XX sequences, which are all used in the exemplification of the present

XX invention. The nucleic acid sequences, protein sequences, and antibodies

XX against them, can be used in the manufacture of a composition. The

XX composition can be used as a medicament (or in the manufacture of a

XX Neisserial bactericidal). For example, some of the identified proteins due to

XX be components of vaccines against Meningococcus B; against all serotypes;

XX and/or against all pathogenic Neisseriae. Identification of sequences

XX from the bacterium will also facilitate production of biological probes,

XX particularly organism-specific probes. Attempts to make efficacious

XX Meningococcus B vaccines have failed mainly due to antigen tolerance.

XX Multivalent vaccines have also been tried but none have successfully

XX overcome antigenic variability. The provision of further, complete

XX sequences may provide an opportunity to identify secreted or surface

XX exposed proteins that may be presumed targets for the immune system and

XX other more variable regions.

Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

## alignment\_scores:

Quality: 238.00 Length: 103  
 Ratio: 3.216 Gaps: 0  
 Percent Similarity: 71.845 Percent Identity: 49.515

## alignment\_block:

US-09-484-577A-4 x AAA81489/rev ..

Align seg 1/1 to reverse of: AAA81489 from: 1 to: 837096

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluTleValG1 36
   ::||| ||| ||||| ||||| ||||| ||||| ||||| |||||
388600 CGCAGCGCGGAGAACAGCGGTTTGGCCCGGCAATTGGAACTGACCGA 388551
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaLal 53
   ::||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
388550 TACGCCGCTCTGCTCCGAAATGGCGCGCGCTTTATTATAGGCGT 388501

```

```

CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
XX Sequence 1305 BP; 303 A; 301 C; 431 G; 270 T; 0 other;
S0
alignment_scores:
    Quality: 235.00      Length: 103
    Ratio: 3.219        Gaps: 0
    Percent Similarity: 70.874      Percent Identity: 49.515
alignment_block:
US-09-484-577A-4 x AAZ54336 ..
Align seg 1/1 to: AAZ54336 from: 1 to: 1305
20 LysThrArgAspGluLeuAlaPheLeuProAlaIleuGluIleValG1 36
   ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
94 CGACACGCGGAGAACAGCGCTTTTGGCCGCGCATTTGGACAGTACC6A 143
   ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
36 uThrProPserProThrAlaArgLeuThrAlaAlaLeuLeuAla1al 53
   ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
144 TAGCGCGGCTCTCGCCGCTCCGAATVGGGCGCGGCTTTATATG6GCT 193
53 euPheTyrCysAlaValAlaTTPAlaGlyLeuGlyArgIleAspIleVal 69
   ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
194 TTTCGCGCTTTGGCTGTGTGTGCTCTGCTGGTGGCAAAATGAAATTTCTG 243
70 AlaSerAlaSerArgIleValProGlyAspArgValLysLeuValG1 86
   ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
244 GCGCGCGCTTCGGCAAAACGGTGTGCGGCGCGCACCAAAACCATCCA 293
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyG1NT 103
   ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
294 GCGGCTCGAAACGGTGTGTTAAAGCGGTAACTGATGCGCGCGGCGACG 343
103 hrValLysAlaGlyIleLeuIleGluLeuAspProPheIleAspGly 119
   ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
344 ATGTGAACAGGAGCAACGCTGCGCAACTGAGAGCTGTGGGAACAGACG 393
120 ValAspVal 122
   ||| ::::: |||
394 AGCGATGTG 402
seq_name: /STDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:AA137329
seq_documentation_block:
ID ID AAT37329 standard; DNA; 28804 BP.
XX
XX AAT37329;
AC
XX
XX 30-NOV-1996 (first entry)
XX
XX Sphingian biosynthetic gene region.
XX
XX Sphingian; polysaccharide; spsb gene; glucosyl-IP-transferase; ds.
XX
XX Sphinomomas strain S88 (ATCC 31554).
XX
XX Key Location/Qualifiers
XX CDS complement (1942..1944)
XX FT /*tag= a
XX FT /codon_start= 1942..1944
XX FT /note= "spsb gene putative initiation codon"
XX FT complement (3311..3313)
XX FT /*tag= b
XX FT /codon_start= 3311..3313

```

```

FT      /note= "spss gene putative initiation codon"
FT      /tag= c
FT      /codon_start= 5323..5325
FT      /note= "spss gene putative initiation codon"
FT      /tag= d
FT      /codon_start= 5526..5528
FT      /note= "spss gene putative initiation codon"
FT      /tag= e
FT      /codon_start= 7076..7078
FT      /note= "spss gene putative initiation codon"
FT      /tag= f
FT      /codon_start= 7588..7590
FT      /note= "spss gene putative initiation codon"
FT      /tag= g
FT      /codon_start= 8643..8645
FT      /note= "spss gene putative initiation codon"
FT      /tag= h
FT      /codon_start= 10938..10940
FT      /note= "spss gene putative initiation codon"
FT      /tag= i
FT      /codon_start= 11569..11571
FT      /note= "spss gene putative initiation codon"
FT      /tag= j
FT      /codon_start= 12886..12888
FT      /note= "spss gene putative initiation codon"
FT      /tag= k
FT      /codon_start= 15165..15167
FT      /note= "spss gene putative initiation codon"
FT      /tag= l
FT      /codon_start= 15883..15885
FT      /note= "spss gene putative initiation codon"
FT      /tag= m
FT      /codon_start= 16748..16750
FT      /note= "spss gene putative initiation codon"
FT      /tag= n
FT      /codon_start= 21082..21084
FT      /note= "spss gene putative initiation codon"
FT      /tag= o
FT      /codon_start= 21706..21708
FT      /note= "spss gene putative initiation codon"
FT      /tag= p
FT      /codon_start= 23238..23240
FT      /note= "spss gene putative initiation codon"
FT      /tag= q
FT      /codon_start= 24113..24115
FT      /note= "spss gene putative initiation codon"
FT      /tag= r
FT      /codon_start= 24683..24685
FT      /note= "spss gene putative initiation codon"
FT      /tag= s
FT      /codon_start= 25744..25746
FT      /note= "spss gene putative initiation codon"
FT      /tag= t
FT      /codon_start= 27534..27536
FT      /note= "spss gene putative initiation codon"

```

```

FT      CDS      27747
FT      /tag= u
FT      /codon_start= 27747..27749
FT      /note= "urif34 gene putative initiation codon"
FT      EP2728841-A2.
FT      28-AUG-1996.
FT      24-JAN-1996; 96EP-0300467.
FT      24-JAN-1995; 95US-0377440.
FT      (SHIN-) SHINETSU BIO INC.
FT      (SHIN-) SHINETSU CHEM CO LTD.
FT      Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;
FT      Yamazaki M;
FT      WPI: 1996-386292/39.
FT      P-PSDB: AAW03997.
FT      New isolated DNA from Sphingomonas sp. - used for transforming
FT      recipient bacteria to obtain hyper-producers of sphing
FT      polysaccharide(s).
FT      Claim 32; Page 56-70; 105pp: English.
FT      A 28.8 kb chromosomal fragment of Sphingomonas strain S88 was
FT      isolated on the basis of its ability to restore sphing
FT      biosynthetic capability to Sphingomonas mutant S88m260. It
FT      contains 23-25 genes, including sps genes coding for biosynthesis of
FT      the polysaccharide sphingon, the genes coding for dTDP-(L)thiamose
FT      biosynthesis, atpB genes coding for a transport function and some
FT      unidentified open translation reading frames (urif). The sps gene
FT      was identified that is believed to code for glucosyl IP-transferase
FT      (AAW03997), an enzyme catalyzing the first step of assembly of
FT      sphingon carboxylates. DNA fragments of 588 can be inserted into
FT      a vector in multiple copies and used to produce engineered bacteria
FT      that are hyper-producers of sphingon.
FT      SQ      Sequence 28804 BP; 4974 A; 9806 C; 9228 G; 4796 T; 0 other:

```

```

alignment_scores:
  Quality: 215.00      Length: 97
  Ratio: 2.905        Gaps: 0
  Percent Similarity: 76.289      Percent Identity: 45.361

```

alignment\_block:

US-09-484-577A-4 x AAW37329/rev ..

Align seg 1/1 to reverse of: AAW37329 from: 1 to: 28804

```

27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAl 43
18808 TTCTGCGCGCGCGCGCTGCGATCATGCAAGCTCGGTCGCCACCCG 18759
43 AArgLeuThrAlaAlaLeuAlaLeuPheTyrCysAlaValAlaT 60
18758 GGGCTCACCGCGCGCGGTGATGCGCGGCTGCGATCACCCGCT 18709
60 TPAlaGlyLeuGlyArgIleAspIleValAlaSerAlaSerArgLysIle 76
18708 GGTGCGCAATGCGCGCGCTGCAATGCGCGCGCGCGCGCGATC 18659
77 ValProGlyAspArgValLysLeuValGluProLeuGluValGlyVal 93
18658 GCCCGATGCGCGCGCGCGATGCGATGCGCGCGCGCGCGCGATC 18609
93 LArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle 110
18608 CCGGCGATTTGTTGTTGCGCGCGCGCGCGCGCGCGCGCGCG 18559

```



Quality: 215.00 Length: 97  
 Ratio: 2.905 Gaps: 0  
 Percent Similarity: 76.289 Percent Identity: 45.361

Alignment block:

US-09-484-577A-4 x AAV99812/rev ..

Align seg 1/1 to reverse of: AAV99812 from: 1 to: 28804

```

27 PheleuproAlaAlaLeuGluIleValGluThrProProserProthral 43
|||||
18808 TTCTGCGCGCGCGCTCGAGATCATCGAACCTCCGCTGCGCCACCGC 18759
43 aargLeuThrAlaAlaLeuLeuAlaLeuPheTyrcysAlaValAla 60
|||||
18758 GCGCCTCACCGCGCGGCTGATGCTGCGCGGCTGCGCATCACACCGCCT 18709
60 rpaLaGlyLeuGlyArgIleAspIleValAlaSerAlaSerArglyIle 76
|||
18708 GCGTGGCAATCGCGCGCGCTGGAAGTGTGCGCGCGCATCGCGCCATC 18659
77 ValProGlyAspArgValLysLeuValGlnProLeuGluValGlyVal 93
|||||
18658 GCCCGCATCGCGCGAGACCAAGATCGTCCGCCGGAAGCGTATCCT 18609
93 lArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyIle 110
|||||
18608 CCGGCGCATTCGTGTGGCGGAGGCGAGAGGTTCAGAGAGCGCGCATG 18559
110 euIlleGluLeuAspProPheAlaGlyGlyValAspValAla 123
|||||
18558 TGATCACGCTCGATCCACCATGTCGCGGCGAGCGCGCCGCCG 18518

```

seq\_name: /SISL1/gcdata/hold-geneSeq/geneSeq-nbml/NA1999.DAT:AAV81474

seq\_documentation\_block:

ID AAV81474 standard; DNA: 28804 BP.

AAV81474:

16-MAR-1999 (first entry)

Chromosomal fragment of *Sphingomonas* strain 88 genome.

Chromosome; sphingane; bacterium; polysaccharide; polymer; additive; food;  
 glycosyl-C55-isoprenylphosphate transferase; textile; cosmetic; paper;  
 paint; cement; viscosity; adhesive; petroleum; chemical; ds.

*Sphingomonas* sp.

Key Location/Qualifiers

FT CDS 21706..23118

FT /tag- a

FT /product- "pspB protein"

PN US854034-A.

PD 29-DEC-1998.

PF 24-JAN-1996; 96US-0592874.

PR 24-JAN-1996; 96US-0592874.

PR 24-JAN-1995; 95US-0377440.

XX (SHIN-) SHIN-ETSU CO INC.

PA (SHIE) SHINETSU CHEM CO LTD.

PI Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;

PI Yamazaki M;

XX

DR MPI: 1999-094809/08.

DR P-PSDR: AAV67750.

XX

PT Production of sphingane polysaccharide products - by introducing DNA  
 from sphingane-producing *Sphingomonas* species in multiple copies into  
 recipient *Sphingomonas* sp.

PS Example 16; Fig 14A-K; 66pp; English.

CC This sequence represents a chromosomal fragment of the *Sphingomonas* sp.  
 CC strain 88 genome. DNA from this sphingane-producing bacterium can be  
 CC used to increase production of sphingane polysaccharides in other  
 CC microorganisms, especially other *Sphingomonas* strains. The DNA  
 CC transferred to other strains includes the pspB gene which encodes a  
 CC glycosyl-C55-isoprenylphosphate (glycosyl-1P) transferase. The sphingane  
 CC are useful as specialty polymers and as additives in textile  
 CC applications, foods, cosmetics, paper, paint, cements, e.g. as viscosity  
 CC modifiers, in various other coating applications, and as adhesives and  
 CC additives to petroleum products and specialty chemicals.

SQ Sequence 28804 BP; 4974 A; 9804 C; 9230 G; 4796 T; 0 other;

Alignment scores:

Quality: 215.00 Length: 97  
 Ratio: 2.905 Gaps: 0  
 Percent Similarity: 76.289 Percent Identity: 45.361

Alignment block:

US-09-484-577A-4 x AAV81474/rev ..

Align seg 1/1 to reverse of: AAV81474 from: 1 to: 28804

```

27 PheleuproAlaAlaLeuGluIleValGluThrProProserProthral 43
|||||
18808 TTCTGCGCGCGCGCTCGAGATCATCGAACCTCCGCTGCGCCACCGC 18759
43 aargLeuThrAlaAlaLeuLeuAlaLeuPheTyrcysAlaValAla 60
|||||
18758 GCGCCTCACCGCGCGGCTGATGCTGCGCGGCTGCGCATCACACCGCCT 18709
60 rpaLaGlyLeuGlyArgIleAspIleValAlaSerAlaSerArglyIle 76
|||
18708 GCGTGGCAATCGCGCGCGCTGGAAGTGTGCGCGCGCATCGCGCCATC 18659
77 ValProGlyAspArgValLysLeuValGlnProLeuGluValGlyVal 93
|||||
18658 GCCCGCATCGCGCGAGACCAAGATCGTCCGCCGGAAGCGTATCCT 18609
93 lArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyIle 110
|||||
18608 CCGGCGCATTCGTGTGGCGGAGGCGAGAGGTTCAGAGAGCGCGCATG 18559
110 euIlleGluLeuAspProPheAlaGlyGlyValAspValAla 123
|||||
18558 TGATCACGCTCGATCCACCATGTCGCGGCGAGCGCGCCGCCG 18518

```

seq\_name: /SISL1/gcdata/hold-geneSeq/geneSeq-nbml/NA1994.DAT:AAQ70050

seq\_documentation\_block:

ID AAQ70050 standard; DNA: 7184 BP.

AAQ70050:

13-MAR-1995 (first entry)

plx gene of *Pasteurella haemolytica*.

*Pasteurella haemolytica*; shipping fever; pasteurellosis; cattle;  
 vaccine; treatment; therapy; prophylaxis; antigen; antibody;  
 immunoassay; ss.

*Pasteurella haemolytica*.

OS

XX

FT

Key Location/Qualifiers

CDS 174..3035

120 Valaspval 122

```

70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuVal
||||:||||| |||:| ||| ||| :|||

```



```

5497 GCGACTGCTCCCGGTAAATTAATTTAGTGGCAGAGTAAGAAATTAA 5546
86 nProleugluValGlyValValAlaThrHisValArgAspGlyGlnT 103
5547 ACCGATTGAAACGCCATTGTACAGAAATTTCTTAAAGATGGCAGT 5596
103 hVallyAlaGlyIleuIleuIleuLeuAspProPheAlaGlyGly 119
5597 TTGTGGAAAGGCGCAATTAATTAAGTCAAGCTTAACGATGGCTTGAT 5646
120 ValAspVal 122
5647 GCGAGATATC 5655

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:AA73218
seq_documentation_block:
ID AA73218 Standard; DNA; 5120 BP.
XX
XX AA73218;
XX
XX 16-FEB-1998 (first entry)
XX
XX DE ApxIBD gene.
XX
XX KW RTX toxin; apxICA gene; apxIBD gene; apxIAB/C gene; apxIABCD gene;
XX repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
XX therapy; A. pleuropneumoniae infection; swine pleuropneumonia; ss.
XX Actinobacillus pleuropneumoniae.
XX
XX FH Key Location/Qualifiers
XX FT -35_signal 524..530
XX FT /*tag= a
XX FT -10_signal 548..553
XX FT /*tag= b
XX FT terminator 572..577
XX FT /*tag= c
XX FT RBS 582..588
XX FT /*tag= d
XX FT CDS 592..2715
XX FT /*tag= e
XX FT /*product= ApxIB
XX FT RBS 2713..2720
XX FT /*tag= f
XX FT CDS 2724..4160
XX FT /*tag= g
XX FT /*product= ApxID
XX
XX CA2170839-A.
XX
XX PD 02-SEP-1996.
XX
XX PF 01-MAR-1996; 96CA-2170839.
XX
XX PR 01-MAR-1995; 95US-0396244.
XX
XX PA (UYGU-) UNIT GUELPH.
XX
XX PI Macinnes J, Mallard B, Ricciattl P, Rosendal S;
XX DR WPI; 1997-245536/23.
XX DR P-PSDB; AAW22153, AAW22154.
XX
XX PT Preparations of microorganisms producing cell-associated RTX toxins
XX PT - especially for production of vaccines against swine
XX PT pleuro-pneumonia
XX
XX PS Disclosure: Pages 97-100; 151pp; English.
XX
XX CC This sequence represents the A. pleuropneumoniae apxIBD gene. This
XX sequence, and the apxICA, apxIAB/C, and apxIABCD genes (see
XX AA73217, AA73219 and AA73220) can be expressed by microorganisms used

```

```

CC in the preparations of the invention. The preparations are bacterial
CC preparations comprising one or more isolated and purified strains of a
CC microorganism that produces one or more RTX (repeat in toxins) toxins,
CC where the strains have at least one cell-associated RTX toxin. The
CC preparations are used for production of vaccines for the prophylaxis and
CC treatment of infectious diseases caused by microorganisms that produce
CC RTX toxins, where the strains have been attenuated or inactivated. The
CC vaccines are preferably against Actinobacillus pleuropneumoniae
CC infection (swine pleuropneumonia). It has been found that A.
CC pleuropneumonia produces significant quantities of cell-associated RTX
CC toxins when cultured under certain conditions, and that the whole-cell
CC protein composition of the cultures corresponds to that of the whole-cell
CC protein profiles obtained from cells recovered at necropsy from the
CC pleural fluid of infected swine. Vaccination with a bacterin prepared
CC from heat-inactivated cultures having significant quantities of
CC cell-associated RTX toxins give significant protection of swine against
CC challenge with homologous strains.
XX
XX S0 Sequence 5120 BP; 1580 A; 880 C; 1139 G; 1521 T; 0 other:

```

```

alignment_scores:
Quality: 169.00 Length: 102
Ratio: 2.315 Gaps: 0
Percent Similarity: 71.569 Percent Identity: 35.294

```

```

alignment_block:
US-09-484-577A-4 x AA73218 ..

```

```

Align seg 1/1 to: AA73218 from: 1 to: 5120

```

```

20 LysThrArgAspGluLeuAlaPheLeuProAlaIleuGluIleValGln 36
2823 CGAGAAAGCATGAAATGATTTTACTGACACTTATGCTGATTTGA 2872
36 hVallyAlaGlyIleuIleuLeuAspProPheAlaGlyGly 119
2873 AACACCGGTGTCAAAAAACCGATGATGCTTATTAATTTATGCGT 2922
53 eupheryCysAlaValAlaThrPalaGlyLeuGlyArgGlyIleVal 69
2923 TCCATTTTGGCATTTAGTTTTCATTTGACATGACGACGATGATG 2972
70 AlaSerAlaSerArgGlyIleValProGlyAspArgValIleuVal 86
2973 GCGACCGCAACGGGTAAATTAATTAAGCGACCGTGAACAAGAAATTAA 3022
86 nProleugluValGlyValValAlaThrHisValArgAspGlyGlnT 103
3023 GCCGATTGAAACGCCCTTGCTTAAAGAAATCTTGTGCAAGACGCAAT 3072
103 hVallyAlaGlyIleuIleuLeuAspProPheAlaGlyGly 119
3073 TTGTTGAAAGATCAAGTGTGTTACCTTGACCGCATTTGGAGCGCAT 3122
120 ValAsp 121
3123 GCGGAT 3128

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA288585
seq_documentation_block:
ID AA288585 standard; DNA; 5120 BP.
XX
XX AC AA288585;
XX
XX DT 05-MAY-2000 (first entry)
XX
XX DE A. pleuropneumoniae DNA encoding clytB and clytD.
XX
XX KW RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
XX anti-inflammatory; antitachytic; antilaborative; treatment; pneumonia;
XX pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;

```

KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;  
 KW urinary infection; peritonitis; meningitis; gastroenteritis;  
 KW passive immunization; clytB; clytD; ds.

XX Actinobacillus pleuropneumoniae.

FT Key location/Qualifiers  
 FT CDS 592..2715

FT /\*tag= a  
 FT /product= "clytB"  
 FT CDS 2724..4160  
 FT /\*tag= b  
 FT /product= "clytD"

XX US6019984-A.

XX 01-FEB-2000.

XX 23-DEC-1996; 96US-0772270.

XX 01-MAR-1995; 95US-0396244.

PA (UXG-) UNIV GUELPH.

XX Mallard B, Rosendal S, MacInnes J, Ricciatti P;

XX WPI: 2000-146864/13.

DR P-PSDB: AAV51407, AAV51408.

XX Bacterial preparation comprising microorganisms which produce a member  
 PT of the Repeats in Toxins (RTX) family, useful for treating swine  
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in  
 PT cattle, and sleepy foal disease -

XX Disclosure: Column 61-66; 96pp; English.

XX This invention describes a novel bacterial preparation (I) which  
 CC comprises one or more isolated and purified strains of a microorganism,  
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or  
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats  
 CC in Toxins), where the strain(s) have at least one RTX toxin which is  
 CC cell-associated. The products of the invention have immunostimulatory,  
 CC antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.  
 CC The bacterial preparation may be used as vaccines for the prophylaxis and  
 CC treatment of infectious diseases caused by strains of microorganisms  
 CC which produce one or more RTX toxins. The infectious diseases are swine  
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;  
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping  
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or  
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,  
 CC polyarthritis and abortion in horses; and urinary infections,  
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations  
 CC may also be used to prepare antibodies which may be used as a means of  
 CC passive immunization. This sequence encodes the Actinobacillus  
 CC pleuropneumoniae clytB and clytD proteins described in the method of  
 CC the invention.

XX Sequence 5120 BP; 1580 A; 880 C; 1139 G; 1521 T; 0 other;

XX alignment\_scores:

XX Quality: 169.00 Length: 102  
 XX Ratio: 2.315 Gaps: 0  
 XX Percent Similarity: 71.569 Percent Identity: 35.294

XX alignment\_block:

XX US-09-484-577A-4 x AAZ88585

XX Align seg 1/1 to: AAZ88585 from: 1 to: 5120

20 LysThrArgAspGluLeuAlaPheLeuProAlaIleuGluIleValG1 36  
 CGAGAAAGATGAAATGAAATTTTACCTGCACACTTAGAGCTGATTGA 2872

seq\_name: /SID51/gcdata/hold-geneseq/geneseq-emb1/NA1997.DAT:AA73220

seq\_documentation\_block:

ID AA73220 standard; DNA; 7721 BP.

XX AA73220;

XX 16-FEB-1998 (first entry)

XX ApxIIIBCD gene.

XX RTX toxin; apxICA gene; apxIIB gene; apxIIAB/C gene; apxIIABCD gene;  
 KW Repeat in toxins toxin; cell-associated RTX toxin; vaccine production;  
 KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia; ss.

XX Actinobacillus pleuropneumoniae.

XX Key location/Qualifiers

XX promoter 152..157

XX TATA\_signal 173..178

XX RBS 195..197

XX CDS 212..733

XX /\*tag= d

XX /\*product= ApxIIIC

XX RBS 727..730

XX /\*tag= e

XX /\*tag= f

XX /\*product= ApxIIIA

XX stem\_loop 3903..3935

XX /\*tag= g

XX RBS 3925..3928

XX /\*tag= h

XX CDS 3943..6078

XX /\*tag= i

XX /\*product= ApxIIIB

XX CDS 6090..7523

XX /\*tag= j

XX /\*product= ApxIIID

XX stem\_loop 7580..7615

XX /\*tag= k

XX CA2170839-A.

XX 02-SEP-1996.



Tue Jul 16 05:46:31 2002

us-09-484-577a-4.rng

Page 13



```

/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-940-086A-53 + 66.00 106.85 290.70 4002.20
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-940-035A-53 + 66.00 106.85 290.70 4002.20
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-231-193A-49 + 66.00 106.81 292.20 4017.20
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-486-273A-49 + 66.00 106.81 292.20 4017.20
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-480-474-49 + 66.00 106.81 292.20 4017.20

seq_name: /cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-592-874-1

seq_documentation_block:
; Sequence 1, Application US/08592874
; Patent No. 5854034
; GENERAL INFORMATION:
; APPLICANT: POLLOCK, THOMAS J.
; APPLICANT: YAMAZAKI, MOTOHIDE
; APPLICANT: THORNE, LINDA
; APPLICANT: MIKOLAJCZAK, MARCIA
; APPLICANT: ARMENTROUT, RICHARD W.
; TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JULES E. GOLDBERG
; STREET: 261 MADISON AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,874
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/377,440
; FILING DATE: 24-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDBERG, JULES E.
; REGISTRATION NUMBER: 24,408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-986-4090
; TELEFAX: 212-818-9479
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FRAGMENT TYPE: N-terminal
US-08-592-874-1

alignment_scores:
Quality: 215.00 Length: 97
Ratio: 2.905 Gaps: 0
Percent Similarity: 76.289 Percent Identity: 45.361

alignment_block:
US-09-484-577A-x US-08-592-874-1/rev ..

Align seg 1/1 to reverse of: US-08-592-874-1 from: 1 to: 28804

27 PhleuPrcAlaAlaIeugCluIleValGluInrProFroSerProThrAl 43
|||||
18808 TTCCTCCGCGCGCGCTCGAGATCARGAAGCGCGGTCTCGCCGCCCGC 18759
43 aArgLeuTrpAlaAlaIeuleuAlaIeuleuAlaIeupheTyrcysAlaValAlaI 60
|||||
18758 GCGCCCTACCGCGCGGTGATGTGGCGGGGCTGGCGGATACACCGCCCT 18709

```

[illegible][illegible]

```

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-096-867-2
seq_documentation block:
Sequence 2, Application US/09096867
Patent No. 6030817
GENERAL INFORMATION:
APPLICANT: Pollock, Thomas J
APPLICANT: Mikolajczak, Marcia
APPLICANT: Yamazaki, Motohide
APPLICANT: Thorne, Linda
APPLICANT: Armentrout, Richard W
TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
FILE REFERENCE: seq list for appl filed from pro. appl
CURRENT APPLICATION NUMBER: US/09/096,867
CURRENT FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 60/049,428
EARLIER FILING DATE: 1997-06-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 28804
TYPE: DNA
ORGANISM: Sphingomonas sp. 588
US-09-096-867-2

```

```

alignment_scores:
    Quality: 215.00      Length: 97
    Ratio: 2.905        Gaps: 0
    Percent Similarity: 76.289      Percent Identity: 45.361

alignment_block:
US-09-484-577A-4 x US-09-096-867-2/rev ..

Align seg 1/1 to reverse of: US-09-096-867-2 from: 1 to: 28804

27 pheleuproalalaaleugluilevalglunthProProserProthral 43
|||||
18808 TTCCTGCGCGCGCGCTCGAGATCATGACAGTCGCGGTGTGCCACCGC 18759
43 aargleuthralalaaleuallalaaleuPherlytCysAlaValaIar 60
|||||
18758 GCGCTCACCGCGCGCGGTGATGTGTGGCGCGCGCGATCACCGCCCT 18709
60 rPalaglyleuaglyarTgIleaspiLevalalaseralaserArglyIle 76
|||||
18708 GCCTGGCATTCGCGCGCGTGGAGTGGTGGCGCGCGACGACGAGCGCATC 18659
77 ValProglYasparYVallylsleuValGlnProleugluValaglyVala 93
|||||
18658 GCCCGGATGGCGAGACACCAAGATGTGTGACATGCCCGGAAAGCGGTATGCT 18609
93 lArgalathrHisValaIargaspelyGlnthVallylsAlaaglyIulel 110
|||||
18608 CCGGGCATTCGTGGTGGCGAGGGGACAGAGGTCATCAGAAAGGCCAGCTGC 18559
110 euilegluneuaspProPhelialaglyGlyValaIaspValaIa 123
|||||
18558 TATATCAGCTGATTCCTCCACCATGTGTGGACGCCGAAGCCGCG 18518

seq_name: /cgn2_6/plodata/2/lna/6A_COMB.seq:US-08-772-270A-6
seq documentation block:
; Sequence 6, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: MacInnes, Janet
;

```

```

APPLICANT: Ricciatti, Paul
APPLICANT: Mallard, Bonnie
APPLICANT: Rosendal, Soren
TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/772,270A
FILING DATE: December 23, 1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Michelle
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 6580-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-6

alignment_scores:
Quality: 169.00 Length: 102
Ratio: 2.315 Gaps: 0
Percent Similarity: 71.569 Percent Identity: 35.294

alignment_block:
US-09-484-577A-4 x US-08-772-270A-6 ..
Align seg 1/1 to: US-08-772-270A-6 from: 1 to: 5120

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2823 CGAGAAAGATGAAATGAATTTTACCTGCACACTTAGAGCTGATTTGA 2872
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
||||| ||| ||||| ||| |||||
2873 AACACCGGTGTCAAAAACGAGATGTGCTATTATTATTAATGCGT 2922
53 eupheTyTCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2923 TCCTATTTTGGCATTTAGTTATTTCATTTGTCAGTCAGCTAGAAATTTG 2972
70 AlaSerAlaSerArgIleValAlaProGlyAspArgValIleValG1 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2973 GCGACCGCAACGGTAAATTAGCGTTTACGACCGTAGCAAGAAATTAA 3022
86 nProLeuGluValGlyValAlaArgAlaThrHisValArgAspGlyGlnT 103
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3023 GCGGATTTGAACCGCTTGTTAAAGAAATCTTTGGCAAGACGACAT 3072
103 hValIysAlaGlyIleLeuIleGluLeuAspProPheAlaGlyIly 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

3073 TTGTTGAAAGATCACTGCTGTACACTTGACCGCATTTGGAGCCGAT 3122
120 ValAsp 121
::: |||
3123 GCGGAT 3128

seq_name: /c9n2_6/ptodata/2/lna/5B.COMB.seq:US-08-488-706-1
seq_documentation_block:
; Sequence 1, Application US/08488706
; Patent No. 5994525
; GENERAL INFORMATION:
; APPLICANT: Kamp, Elbarie M.
; APPLICANT: Smits, Marinus A.
; TITLE OF INVENTION: Recombinant Vaccine For Prevention and/or
; TITLE OF INVENTION: Treatment of Pleuropneumonia Infections
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Abelman, Frayne & Schwab
; STREET: 708 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM PC or compatible
; OPERATING SYSTEM: DOS 3.31
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,706
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/722,971
; FILING DATE: 28-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Cinnamon, Jay S.
; REGISTRATION NUMBER: 24,156
; REFERENCE/DOCKET NUMBER: 201,875
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-949-9022
; TELEFAX: 212-949-9190
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8370
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-488-706-1

alignment_scores:
Quality: 169.00 Length: 102
Ratio: 2.315 Gaps: 0
Percent Similarity: 71.569 Percent Identity: 35.294

alignment_block:
US-09-484-577A-4 x US-08-488-706-1 ..
Align seg 1/1 to: US-08-488-706-1 from: 1 to: 8370

20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluIleValG1 36
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6073 CGAGAAAGATGAAATGAATTTTACCTGCACACTTAGAGCTGATTTGA 6122
36 uThrProSerProThrAlaArgLeuThrAlaAlaLeuAlaAlaL 53
||||| ||| ||||| ||| |||||
6123 AACACCGGTGTCAAAAACGAGATGTGCTATTATTATTAATGCGT 6172
53 eupheTyTCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIleVal 69
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6173 TCCTATTTTGGCATTTAGTTATTTCATTTGTCAGTCAGCTAGAAATTTG 6222

```



70 AlaseraLaserArgLysIleValProGlyAspArgValLysLeuValG1 86  
 6223 GCGACCGCAACGGGTAAATTAGCTTTAGCGACCGGTAGCAAGAAATTAA 6272  
 86 nProLeuGluValGlyValAlaIArgAlaIThrHisValAlaIArgAspGlyLnt 103  
 6273 GCCCATTCGAAACGCGTGGTTAAAGAAATCTTGTGCAAGACGCAAT 6322  
 103 hrValLysAlaGlyLnuIleuIleGluLeuAspProphealaglyGly 119  
 6323 TTGTGTGAAGAAGATCAGTCTGCTTTACACTTGACCGCATTTGGAGCGCAT 6372  
 120 Valasp 121  
 6373 GCGGAT 6378

seq\_name: /cgn2\_6/ptodata/2/lna/6A\_COMB.seq:US-08-772-270A-14

documentation\_block:  
 Sequence 14, Application US/08772270A  
 Patent No. 6019984  
 GENERAL INFORMATION:  
 APPLICANT: Macines, Janet  
 APPLICANT: Ricciatti, Paul  
 APPLICANT: Mallard, Bonnie  
 APPLICANT: Rosendal, Soren  
 TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR  
 TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Bereskin & Parr  
 STREET: 40 King Street West  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5H 3Y2  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/772,270A  
 FILING DATE: December 23, 1996  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gravelle, Micheline  
 REGISTRATION NUMBER: 40,261  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7721 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 ORGANISM: Actinobacillus pleuropneumoniae  
 US-08-772-270A-14

alignment\_scores:  
 Quality: 168.00 Length: 102  
 Ratio: 2.333 Gaps: 0  
 Percent Similarity: 70.588 Percent Identity: 35.294

alignment\_block:  
 US-09-484-577A-4 x US-08-772-270A-14 ..

Align seg 1/1 to: US-08-772-270A-14 from: 1 to: 7721

12 ArgGluValAlaValAlaGlyLysThrArgAspGluLeuAlaPhe 28  
 6165 CGCAACATTTAGATATACCCACGACACAAAGATGAAAGCATTTT 6214  
 28 uProAlaAlaLeuGluIleValGluThrProPheProThrAlaArgL 45  
 6215 GCCTCGCATTTAGAGTTATATGACACCTATTTCAAAAAGCCACGCG 6264  
 45 euThrAlaAlaLeuAlaAlaLeuPheTyrCysAlaValAlaTrpAla 61  
 6265 TGATCGCTATTGTGATATGATGCTATTGTTTATTTAGCTATTGTAATTC 6314  
 62 GlyLeuGlyArgIleAspIleValAlaLaserAlaLaserArgLysIleValPr 78  
 6315 ATTATTAGTAAGTAGAATTTGCTGCTGCTACAGTAAGTTGTTT 6364  
 78 oGlyAspArgValLysLeuValGlnProLeuGluValGlyValAlaArg 95  
 6365 TAGTGACATAGTAAAGAAATTAAGCTATTGACATGCTTAGTAAAG 6414  
 95 laThrHisValAlaArgAspGlyGlnThrValLysAlaGlyLnuIleuLe 111  
 6415 ACATTTTGTTAAGATGAGCAATTGTTGAAAAAGCAAAATTATTATTA 6464  
 112 GluLeu 113  
 6465 AATCTC 6470

seq\_name: /cgn2\_6/ptodata/2/lna/6B\_COMB.seq:US-08-976-259-79

seq\_documentation\_block:  
 Sequence 79, Application US/08976259  
 Patent No. 6316609  
 GENERAL INFORMATION:  
 APPLICANT: Dillon, Patrick J.  
 APPLICANT: Choi, Gil H.  
 APPLICANT: Welch, Rodney A.  
 TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli  
 Patent No. 6316609  
 NUMBER OF SEQUENCES: 142  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 STREET: 1100 New York Ave, N.W., Suite 600  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005-3934  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/976,259  
 FILING DATE: Herewith  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/031,626 AND US 60/061,953  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Steffe, Eric K.  
 REGISTRATION NUMBER: 36,688  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2500  
 TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 79:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3576 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear



```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)-205-8000
TELEFAX: (703)-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
SPRAIN: Serratia marcescens Str41
US-09-005-232A-1

```

```

alignment_scores:
Quality: 99.00      Length: 57
Ratio: 2.538        Gaps: 2
Percent Similarity: 68.421  Percent Identity: 43.860

```

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-491-772-5

```

```

Align seg 1/1 to: US-09-005-232A-1 from: 1 to: 4547
58 ValAlaTrpAlaGlyLeuGlyArgGlyLeuAspIle...ValAlaSerAlaase 73
:::|||||:::|||||:::|||||:::|||||:::
1936 TTACCTGGGGGGGCTTGGCCCTTGATAAAGGGGTGGCGCGGG 1985
73 rArgLysIleValProGlyAspArgValLysLeuValGlnProLeuGluV 90
||| |||:::||||| ||| |||
1986 ATCGTAAACCGCTTCCGCAACCGC...AAACGGTGCAGGCCCGGCCCA 2032
90 alGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAla 106
|||||:::|||||:::|||||:::|||||
2033 CGCGCATCATTAAGAAATATTCGGCTGACAGATGGCGCAAAAGTGAAGCC 2082
107 GlyLysIleLeuIleGluLeu 113
|||||:::|||||:::|||||
2083 GGTGAGTCTGCTGTCAGCTC 2103

```

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-491-772-5
seq_documentation_block:
; Sequence 5, Application US/09491772
; Patent No. 6329172
; GENERAL INFORMATION:
; APPLICANT: Rhee, Joon Shick
; APPLICANT: Pan, Jae Gu
; APPLICANT: Ahn, Jung Hoon
; TITLE OF INVENTION: ABC Transporter Gene Cluster in Pseudomonas fluorescens for
; FILE REFERENCE: Enhanced Lipase Secretion
; CURRENT APPLICATION NUMBER: US/09/491,772
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: KR 10-1999-0008512
; NUMBER OF SEQ ID NOS: 1999-03-13
; SOFTWARE: KOPATIN 1.5
; SEQ ID NO 5
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Pseudomonas fluorescens
US-09-491-772-5

```

```

alignment_scores:
Quality: 90.00      Length: 63
Ratio: 2.250        Gaps: 0
Percent Similarity: 63.492  Percent Identity: 36.508

```

```

alignment_block:

```

```

US-09-484-577A-4 x US-09-491-772-5
Align seg 1/1 to: US-09-491-772-5 from: 1 to: 1302

```

```

60 TrpAlaGlyLeuGlyArgGlyLeuValAlaSerAlaSerArgLysIle 76
|||||:::|||||:::|||||:::|||||:::
94 TGGCCAGCGCTGGCGCCGTCGACCAAGGCATTCGGTGCAGGACCGCT 143
76 eValProGlyAspArgValLysLeuValGlnProLeuGluValGlyVal 93
||| ::| ||| ||||| ::::| |||||
144 CGTGTCTGGGCAGCGCAAGCGGTGCAAACTTCAGCCCGGCGGTGG 193
93 aArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyLysIle 109
||| |||||:::|||||:::|||||
194 TCAGCCGATTCGTGGTGGCGGCGAAGCGTTAAACAAGCGACCGG 243
244 CTGTTTCGCTGACCAAGCCAGACCGAGCTGATGTG 282

```

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-491-772-1

```

```

seq_documentation_block:
; Sequence 1, Application US/09491772
; Patent No. 6329172
; GENERAL INFORMATION:
; APPLICANT: Rhee, Joon Shick
; APPLICANT: Pan, Jae Gu
; APPLICANT: Ahn, Jung Hoon
; TITLE OF INVENTION: ABC Transporter Gene Cluster in Pseudomonas fluorescens fo
; FILE REFERENCE: Enhanced Lipase Secretion
; CURRENT APPLICATION NUMBER: US/09/491,772
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: KR 10-1999-0008512
; NUMBER OF SEQ ID NOS: 1999-03-13
; SOFTWARE: KOPATIN 1.5
; SEQ ID NO 1
; LENGTH: 8580
; TYPE: DNA
; ORGANISM: Pseudomonas fluorescens
US-09-491-772-1

```

```

alignment_scores:
Quality: 90.00      Length: 63
Ratio: 2.250        Gaps: 0
Percent Similarity: 63.492  Percent Identity: 36.508

```

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-491-772-1

```

```

Align seg 1/1 to: US-09-491-772-1 from: 1 to: 8580
60 TrpAlaGlyLeuGlyArgGlyLeuValAlaSerAlaSerArgLysIle 76
|||||:::|||||:::|||||:::|||||:::
4085 TGGCCAGCGCTGGCGCCGTCGACCAAGGCATTCGGTGCAGGACCGCT 4134
76 eValProGlyAspArgValLysLeuValGlnProLeuGluValGlyVal 93
||| ::| ||| ||||| ::::| |||||
4135 CGTGTCTGGGCAGCGCAAGCGGTGCAAACTTCAGCCCGGCGGTGG 4184
93 aArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyLysIle 109
||| |||||:::|||||:::|||||
4185 TCAGCCGATTCGTGGTGGCGGCGAAGCGTTAAACAAGCGACCGG 4234
110 LeuIleGluLeuAspProPheAlaGlyLysValAspVal 122
|||||:::|||||:::|||||:::|||||
4235 CTGTTTCGCTGACCAAGCCAGACCGAGCTGATGTG 4273

```

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-614-770A-1

```

```

Alignment_scores:
      Quality: 77.00      Length: 108
      Ratio: 1.453      Gaps: 4
      Percent Similarity: 49.074      Percent Identity: 29.630

Alignment_block:
US-09-484-577A-4 x US-08-614-770A-1 ..

Align seg 1/1 to: US-08-614-770A-1 from: 1 to: 49272

14 ValValaIaValaIGlyLyThrRAspGluLeuAlaPheLeuProAl 30
||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
41841 GIGATGCCGTTGAAACACACGATCCGGCGCGGATCTCGAAATTGTCCAC 41890

30 aAlaLeu.....GluIleValaGluT 37
|||||T|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

41891 GCCCTTCCTCAGGTTCTCTGAGGCCACGTCGGCGTCTGGAGATCACCCGA 41940

37 hrProPterProthrAlaArgLeuThrAlaIleAlaLeuAlaLeu 53
||| ||| |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

41941 CCAAcCCAGACGGCGGCGAGGGCGATGCGTGGCGGATGCGCTTC 41990

54 pheTYrYSaIaValaAlaTrp.....AlaGlyLeuGlyArgIleAspI 68
||| ||| |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

41991 TTCATGTGCGGTACTTCTTGTTGGTGGCTGTCAAGTACGATGATCAAGT 42040

```

seq\_name: /can2\_6/ptodata/

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-188-582-15

```

seq_documentation_block:
; Sequence 15, Application US/08188582
; Patent No. 5534410
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3603 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2214
US-08-188-582-15

alignment_scores:
Quality: 74.50 Length: 117
Ratio: 1.202 Gaps: 4
Percent Similarity: 52.991 Percent Identity: 27.350

alignment_block:
US-09-484-577a-4 x US-08-188-582-15 ..
Align seg 1/1 to: US-08-188-582-15 from: 1 to: 3603

4 LeuserProArqmetlysserAlaargGluValAlaValGlyGly 20
111 111 111 111 111 111 111 111 111 111
913 CTGTGTCCTTCTCGAG.....AGAGCTTACCGCGCTTGAGACAGCT 956
20 sThArqAspGluLeuAlaPheLeuProAlaAlaLeuGluLeuAla 37
111 111 111 111 111 111 111 111 111 111
957 GAACCGCGACTCGCGGCTTATCCAG.....CAGAGCCAGCAGC 997
37 hrProPserProThrAlaArgLeuThrAlaAlaLeuAlaAlaLeu 53
111 111 111 111 111 111 111 111 111 111

```

```

seq_name: /cgn2.6/ptodata/2/1na/5A_COMB.seq:US-08-646-715-15
1212 G 1212

seq_documentation_block:
; Sequence 15, Application US/08646715
; Patent No. 5637686
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3603 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA

```

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2214  
US-08-646-715-15

## alignment\_scores:

Quality: 74.50 Length: 117  
Ratio: 1.202 Gaps: 4  
Percent Similarity: 52.991 Percent Identity: 27.350

## alignment\_block:

US-09-484-577a-4 x US-08-646-715-15 ..

Align seg 1/1 to: US-08-646-715-15 from: 1 to: 3603

```
4 leuserProArgMetIysSerAlaArgGluValAlaValAlaGlyGly 20
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
913 CTTGTGCTTCTTCTGAG.....AGGAGCTTACCCGCTTGAGACAGCT 956
20 sthrArgaspGluLeuAlaPheLeuProAlaAlaLeuGluIleValGluT 37
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
957 GACCCCGACTCCGCGCTTCATCCAG.....CAGAGCCAGCAGC 997
37 hrProProSerProThrAlaArgLeuThrAlaAlaLeuAlaLeu 53
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
998 AGCCGCCACCGCCACCTGCAGCGCAGCAGCTGCGCTCAGCGCGTG 1047
54 PheTyrCysAlaValAlaTPrAlaGlyLeuGlyArgIleAspIleValAl 70
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1048 CTGAGTAGCTCGGTCCAGCGCGCCGCGAGAGCGCGGCCACCGTGAC 1097
70 aSerAlaSerArgIleValProGlyAspArgValIysLeuValGlnP 87
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1098 CAGTCCCTCCAGCCCCCTGTG.....CTCAGCCTCAGCGCAGC 1135
87 roLeuGluValAlaValAlaArgAlaThrHisValAlaArgaspGlyInThr 103
|| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1136 CCAGCAGGTGGCGTC.....GGCAGCAG 1161
104 ValIysAlaGlyGluIleLeuIleGluLeuAspProPheAlaGlyVal 120
::: :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1162 GGGCAACCCACACCGCTGTATCCAGCAGCTCCGAGAGCCAGAGCCCT 1211
120 1 120
1212 G 1212
```

